QUANTUM ENTANGLEMENT FOR
SYSTEMS OF IDENTICAL BOSONS
I. GENERAL FEATURES

B. J. Dalton[1,2], J. Goold[1,3], B. M. Garraway[4] and M. D. Reid[2]

[1] *Physics Department, University College Cork, Cork City, Ireland*
[2] *Centre for Quantum and Optical Science*, Swinburne University of Technology, Melbourne, Victoria 3122, Australia*
[3] *The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste 34151, Italy*
[4] *Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN19QH, United Kingdom*

Email: bdalton@swin.edu.au
∗ Formerly *Centre for Atom Optics and Ultrafast Spectroscopy*

Abstract

These two accompanying papers are concerned with two mode entanglement for systems of identical massive bosons and the relationship to spin squeezing and other quantum correlation effects. Entanglement is a key quantum feature of composite systems in which the probabilities for joint measurements on the composite sub-systems are no longer determined from measurement probabilities on the separate sub-systems. There are many aspects of entanglement that can be studied. This two-part review focuses on the meaning of entanglement, the quantum paradoxes associated with entangled states, and the important tests that allow an experimentalist to determine whether a quantum state - in particular, one for massive bosons is entangled. An overall outcome of the review is to distinguish criteria (and hence experiments) for entanglement that fully utilise the symmetrisation principle and the super-selection rules that can be applied to bosonic massive particles.

In the first paper (**I**), the background is given for the meaning of entanglement in the context of systems of identical particles. For such systems, the requirement is that the relevant quantum density operators must satisfy the symmetrisation principle and that global and local super-selection rules prohibit states in which there are coherences between differing particle numbers. The justification for these requirements is fully discussed. In the second quantisation approach that is used, both the system and the sub-systems are modes (or sets of modes) rather than particles, particles being associated with different occupancies of the modes. The definition of entangled states is based on first defining the non-entangled states - after specifying which modes constitute the sub-systems. This work mainly focuses on the two mode entanglement for massive bosons, but is put in the context of tests of local hidden variable theories, where one may not be able to make the above restrictions. The review provides the detailed arguments necessary for the conclusions of a recent paper, where the question of how to rigorously demonstrate the entanglement of a two-mode Bose-Einstein condensate (BEC) has been examined.

In the accompanying review paper (**II**), we consider spin squeezing and other tests for entanglement that have been proposed for two-mode bosonic systems. We apply the approach of review (I) to determine which tests, and which modifications of the tests, are useful for detecting entanglement in massive bosonic (BEC), as opposed to photonic, systems. Several new inequalities are derived, a theory for the required two-mode interferometry is presented, and key experiments to date are analysed.

PACS Numbers   03.65 Ud, 03.67 Bg, 03.67 Mn, 03.75 Gg

Contents

1 Introduction

Since the paradoxes of Einstein-Podolski-Rosen (EPR) [1], Schrodinger [2] and Bell [3], *entanglement* has been recognised as a key feature that distinguishes quantum physics from classical physics. Entangled quantum states underlie the EPR and Bell paradoxes, which reveal the conflict between quantum mechanics and local realism, and the famous Schrodinger cat paradox, where a cat is apparently prepared in a state simultaneously both dead and alive. Entanglement not only provides a way to rigorously test quantum principles, but is the basis for the many quantum information tasks like quantum cryptography. Despite the fundamental interest, there have been only a few experimental tests of entanglement for systems of massive particles. Yet, the substantial recent progress in cooling atomic systems, in particular to form Bose-Einstein condensates, makes such entanglement tests much more feasible.

In this review (**I**), we explain the meaning of entanglement, and examine how to verify entanglement, for systems of identical boson particles. This leads us to focus on symmetrisation and superselection rules, and to consider their implication for entanglement criteria when applied to massive bosonic particles. This paper provides the theoretical background for a recent paper [4] and a subsequent review (**II**) that analyses the suitability of specific criteria, new and old, to detect entanglement in bosonic systems, and applies the criteria to interpret experimental findings.

As well as reviewing the topic and presenting some new results in review II, these two articles are intended as comprehensive papers for post docs and postgraduates who are changing field or starting work in a new one and need to gain a thorough understanding of the present state of knowledge. With this aim in mind we have not followed the conventional approach in review articles of merely quoting formulae and giving references, but instead have presented full proofs of the key results. To really understand a field, we believe it is necessary to work through the derivations. However, in order to shorten the main body of the articles, we have included many of the details in Appendices.

1.1 Entanglement: Definitions and Historical Context

Entanglement arises in the context of *composite* quantum systems composed of distinct components or *sub-systems* and is distinct from other features of quantum physics such as *quantization* for measured values of physical quantities, *probabilistic outcomes* for such measurements, *uncertainty principles* involving pairs of physical quantities and so on. Such sub-systems are usually associated with sub-sets of the physical quantities applying to the overall system, and in general more than one choice of sub-systems can be made. The formalism of quantum theory treats *pure states* for systems made up of two or more distinct sub-systems via tensor products of sub-system states, and since these product states exist in a Hilbert space, it follows that linear combinations of such products could also represent possible pure quantum states for the system. Such *quantum superpositions* which cannot be expressed as a *single* product of

sub-system states are known as *entangled* (or *non-separable*) states.

The concept of entanglement can then be extended to *mixed states*, where quantum states for the system and the sub-systems are specified by density operators. The detailed definition of entangled states is set out in Section 2. This definition is based on first carefully defining the *non-entangled* (or *separable*) states. The set of non-entangled states must allow *all* possible quantum states for the given sub-system, but in addition these states must be *preparable* via processes involving *separate operations* on each sub-system after which correlated sub-system quantum states are combined in accordance with *classical probabilities*. Thus, although the sub-system states retain their quantum natures the combination resulting in the overall system state is formed classically rather than quantum mechanically. This overall process then involves *local operations* on the distinct sub-systems and *classical communication* (*LOCC*) to prepare a general non-entangled state. The entangled states are then just the quantum states which are *not* non-entangled states. The general idea that in all composite systems the non-entangled states all involve LOCC preparation processes was first suggested by Werner [5]. The notion of *quantum states*, the nature of the systems and *sub-systems* involved and the specific features required in the definition of non-entangled states when *identical particles* are involved is discussed in detail in Section 3. Entangled states underlie a number of effects that cannot be interpreted in terms of classical physics, including *spin squeezing*, non-local measurement *correlations* - such as for the Einstein-Podolski-Rosen *(EPR) paradox* and violations of *Bell Inequalities* ([1], [2], [3], [6], [7]). The quantum theory of measurement [8], [9], [10], [11] invokes entangled states of the system and measuring apparatus as key concepts in the theory. More recently, entangled states have been recognised as a resource that can be used in various *quantum technologies* for applications such as teleportation, quantum cryptography, quantum computing and so on. Recent expositions on the effects of entanglement and its role in *quantum information science* include [11], [12], [13], [14], [15], [16].

It would be pointless to characterise states as entangled unless such states have some important properties. The key requirement is that entangled states exhibit a novel *quantum feature* that is only found in *composite* systems. As will be seen in SubSection 2.3, *separable* states are such that the *joint probability* for measurements of all physical quantities associated with the sub-systems can be found from separate measurement probabilities obtained from the sub-system density operators and the overall classical probability for creating particular products of sub-system states. In general, entangled states do not exhibit this feature of separable probabilities, and it is this key non-separability feature that led Schrodinger to call these states "entangled". Where the sub-systems are spatially separated, one can define spacelike separated local measurement events on each. This was historically the reason why the sub-systems and their measurements are often referred to as *local*. The EPR paper [1] suggested the possibility that although the predictions of quantum theory were correct, the theory was *incomplete* and there was an underlying reality in the form of *classical hidden variables*. Averaging over the unknown values of the hidden variables

7

would be required to produce the same measurement probabilities as quantum theory. *Local hidden variable* theories (LHV) are discussed in SubSection 2.5, and it will be seen that the joint probabilities for measurements of sub-system physical quantities are of the same form as for *separable* states. As will be seen in SubSections 2.6 and 2.7, EPR or Bell inequality violations do *not* occur for states described either by LHV theories or as quantum separable states. Hence there is a direct link between EPR and Bell violations and both the *failure* of LHV theories and the presence of *entanglement*. The fact that certain entangled states do not exhibit the feature of separable probabilities shown in classical LHV theories highlights entanglement being a *non-classical* feature found in composite systems.

Note that although an EPR or Bell inequality violation requires the quantum state to be entangled, there are examples of mixed entangled states that do *not* violate a Bell inequality. For *pure* states of qubits Gisin [17] showed that entangled states always violated Bell inequality, but for *mixed* states Werner [5] and others [18], [19], [20] have shown there are entangled states (*Werner states* [13]) for which a hidden variable theory can be constructed that gives the same joint probability function for measurement outcomes as quantum theory. Of course the quantum measurement outcomes must be given before the hidden variable model is constructed - there is no known way to determine the LHV theory distribution functions independently. These specific entangled states will therefore satisfy Bell inequalities.

The mixed entangled states considered by Werner [5] for which a hidden variable theory could be constructed were of a special form. Two distinguishable sub-systems each with d basis states $|u_r\rangle$ were considered, for which the states could be transformed by unitary operators $\widehat{U}$, and the combined density operator $\widehat{\rho}$ was required to be invariant under all unitary transformations of the form $\widehat{U} \otimes \widehat{U}$, so that $\widehat{\rho} = (\widehat{U} \otimes \widehat{U})\widehat{\rho}(\widehat{U}^\dagger \otimes \widehat{U}^\dagger)$. Werner [5] considered the following unitary operators: (a) $\widehat{U}_{-r}$ such that $\widehat{U}_{-r}|u_r\rangle = -|u_r\rangle$, $\widehat{U}_{-r}|u_s\rangle = +|u_s\rangle$ for $s \neq r$ (b) $\widehat{P}(r \to \mu r)$, which permute basis states $\widehat{P}(r \to \mu r)|u_r\rangle = |u_{\mu r}\rangle$ (c) $\widehat{U}_{rot}(n,m)$, which transform basis states $|u_n\rangle$, $|u_m\rangle$ into linear combinations of each other $\widehat{U}_{rot}(n,m)|u_n\rangle = U_{nn}|u_n\rangle + U_{nm}|u_m\rangle$, $\widehat{U}_{rot}(n,m)|u_m\rangle = U_{mn}|u_n\rangle + U_{mm}|u_m\rangle$. As a consequence of these invariances Werner [5] showed that the density operator could be expressed in terms of a single parameter Φ in the form $\widehat{\rho} = (d^3 - d)^{-1}\left\{(d - \Phi)\widehat{1} + (d\Phi - 1)\widehat{V}\right\}$, where $\widehat{1}$ is the unit operator and $\widehat{V}$ is the flip operator. These have matrix elements $\left(\widehat{1}\right)_{rs,\,nm} = \delta_{rn}\,\delta_{sm}$ and $\left(\widehat{V}\right)_{rs,\,nm} = \delta_{rm}\,\delta_{sn}$. From this form of the density operator Werner [5] showed that the probability function for joint measurement outcomes on the two sub-systems could be expressed in the same form as applied in hidden variable theory. So although the mixed entangled state Werner considered were of a restricted type, the work demonstrated that entanglement did not preclude all hidden variable theory interpretations of the joint measurements. The fact that some entangled states do not violate a Bell inequality is another consequence of

Werner's result.

The issue of how best to treat the quantum aspects of *correlations* in measurement outcomes in composite quantum systems is still an active area of research and is beyond the scope of these two papers. Quantum entanglement is clearly relevant to the discussion, but concepts such as *quantum discord* [21], [22] and *EPR steering* [23], [24], [25] are now being used to describe quantum correlations. The link between these concepts is discussed in [26]. In these recent discussions of quantum correlation, it turns out that some separable states are regarded as exhibiting quantum correlations.

It is now generally recognised that entanglement is a *relative* concept ([27], [28], [29]), [13], [30], [31] and not only depends on the quantum state under discussion but also on which *sub-systems* are being considered as entangled (or non-entangled). A quantum state may be entangled for one choice of the subsystems but may be non-entangled if another choice of sub-systems is made. A simple example often cited is that for the hydrogen atom [29], a system made up of two distinguishable particles, a proton and an electron. Here the energy eigenstates are non-entangled if the sub-systems refer to the centre of mass of the entire atom and the relative position of the electron and the proton, but which would be entangled if the sub-systems were the positions of the electron and proton. It could be argued that the centre of mass and the relative position are not really independent sub-systems - one always accompanies the other - but as unrelated centre of mass and relative position quantum states can be prepared, they can be regarded as distinct sub-systems. The individual positions of the electron and the proton are also distinct sub-systems, and the ground state of the hydrogen atom is indeed entangled - the electron position is tightly correlated with the proton position. Another example involves two different choices of single particle states in a two mode Bose-Einstein condensate (BEC) - a system with a large number of identical particles. The issue of defining sub-systems will be dealt with below, but taking the original two sub-systems to be bosonic modes (or single particle states) denoted $|\phi_A\rangle$ and $|\phi_B\rangle$, a well known N boson entangled state of these two modes A and B (with mode annihilation operators $\widehat{a}$ and $\widehat{b}$) is the *binomial state* given by $|\Phi\rangle = ((\cos\theta \exp(i\chi/2)\widehat{a} + \sin\theta \exp(-i\chi/2)\widehat{b})^{\dagger})^N/\sqrt{N!}\ |0\rangle$ (see [32] and Paper II, **Section 3.7**) which is a quantum superposition of Fock states $(\widehat{a}^{\dagger})^k/\sqrt{k!}\ (\widehat{b}^{\dagger})^{N-k}/\sqrt{(N-k)!}\ |0\rangle$ with $k = 0, ..., N$. Introducing new modes via $\widehat{c} = (\cos\theta \exp(i\chi/2)\widehat{a} + \sin\theta \exp(-i\chi/2)\widehat{b})$ and $\widehat{d} = (-\sin\theta \exp(i\chi/2)\widehat{a} + \cos\theta \exp(-i\chi/2)\widehat{b})$ we see that we can also write $|\Phi\rangle = (\widehat{c}^{\dagger})^N/\sqrt{N!}\ |0\rangle$, so that the same quantum state is a separable state if the sub-systems are chosen to be the new modes C and D. Another example is the ground state of the single mode non-interacting BEC trapped in a harmonic oscillator (HO) potential. This is a separable state, with all bosons in the lowest energy mode if the sub-systems are chosen as the HO modes. However, if **single** particle position states spatially localised in two different regions are chosen as two sub-systems, then the same ground state for the identical particle system is spatially entangled, as pointed out by Goold et al [33].

1.2 Measures and Tests for Entanglement

Various *measures* of entanglement have been defined for certain types of quantum state - see [13], [14], [30], [31], [34], [22], [35], for details of these, and are aimed at quantifying entanglement to determine which states are more entangled than others. This is important since entanglement is considered as a resource needed in various quantum technologies. Calculations based on such measures of entanglement confirm that for some choices of sub-systems the quantum state is entangled, for others it is non-entangled. For two mode pure states the *entanglement entropy* - being the difference between the entropy for the pure state (zero) and that associated with the reduced density operator for either of the two sub-systems - is a useful entanglement measure. As entropy and information changes are directly linked [13], [14], this measure is of importance to *quantum information science*. Measurements of entanglement based on *Renyi entropy* and *purity* are discussed in [36], [37] and [38]. Another entanglement measure is *particle entanglement*, defined by Wiseman et al [39], [40], [34] for identical particle systems and based on projecting the quantum state onto states with definite particle numbers. One of the problems with entanglement measures is that there is often no simple way to measure the quantities required.

In the case of *bipartite* entanglement in qubit systems [41] and [42] obtained a sufficient condition for a quantum state to be entangled (*PPT condition*) (see [31], [30] for details). Suppose the density operator $\widehat{\rho}$ is changed into $\widehat{\rho}^T$ by mapping the matrix elements associated with one of the sub-systems into their *transpose*. Then provided the new operator $\widehat{\rho}^T$ is a valid density operator (with real, non-negative eigenvalues that add to unity), the original density operator represents a separable state. Thus, if it is shown that some of the eigenvalues of $\widehat{\rho}^T$ are negative, then the state $\widehat{\rho}$ is entangled. However, it is often not practical to use this as an entanglement test for systems with large numbers of basis states, as it requires being able to measure all the density matrix elements. It was also later realised [43] that in general, the PPT condition was not a necessary condition for entanglement, apart from cases of 2×2 and 2×3 subsystems - that is, showing that $\widehat{\rho}^T$ has only positive eigenvalues will not guarantee that $\widehat{\rho}$ is separable, as counter-examples for 2×4 and 3×3 subsystems showed.

Although not directly related to the various quantitative measures of entanglement, the results for certain measurements can play the role of being *signatures* or *witnesses* or *tests* of entanglement [30], [31], [34]. These are in the form of *inequalities* for *variances* and *mean values* for certain physical quantities, which are dependent on the inequalities applying for non-entangled quantum states. If such inequalities are *violated* then it can be concluded that the state is *entangled* for the relevant sub-systems. In the case of entanglement witnesses, the idea is to find a hermitian operator $\widehat{W}$ such that for separable states $Tr(\widehat{W}\widehat{\rho}) \geq 0$, so that if $Tr(\widehat{W}\widehat{\rho}) < 0$ the state must be entangled. Here we note that the density operator occurs linearly when evaluating the quantities involved. Some of the *correlation* tests discussed in paper II are cases involving entanglement witnesses. However, in more general tests for entanglement the density operator appears non-linearly. For example, a *spin squeezing*

test for entanglement may require showing that the variance for a spin operator is less than a multiple of the magnitude of the mean value of another spin operator - thus for example $\left\langle \Delta \widehat{S}_x^2 \right\rangle < |\left\langle \widehat{S}_z \right\rangle|/2$. This could be written as $Tr((\widehat{S}_x^2 \pm \widehat{S}_z/2)\widehat{\rho}) - (Tr(\widehat{S}_x)\widehat{\rho}))^2 < 0$, which is of a more general form than for an entanglement witness. Non-linear tests are discussed in Ref. [31]. One of the advantages of entanglement tests is that the quantities involved can be measured. It cannot be emphasised enough that these tests provide *sufficiency conditions* for establishing that a state is entangled. So if the test is satisfied we can conclude that the state is not separable. The failure of a test does *not* mean that the state is not entangled - sufficiency does not imply *necessity*. The violation of a *Bell inequality* is an example of such a signature of entanglement, and the demonstration of *spin squeezing* is regarded as another. However, the absence of spin squeezing (for example) does not guarantee non-entanglement, as the case of the *NOON* state in SubSection **3.6** of the accompanying paper II shows. A significant number of such inequalities have now been proposed and such signatures of entanglement are the primary focus of the accompanying paper, which is aimed at identifying which of these inequalities really do identify entangled states, especially in the context of *two mode* systems of *identical bosons*.

At present there is *no clear linkage* between quantitative measures of entanglement (such as entanglement entropy) and the quantities used in conjunction with the various entanglement tests (such as the relative spin fluctuation in spin squeezing experiments). Results for experiments demonstrating such non-classical effects cannot yet be used to say much more than the state *is* entangled, whereas ideally these experiments should determine *how* entangled the state is. Again we emphasise that neither the entanglement tests nor the entanglement measures are being used to *define* entanglement. Entanglement is defined first as being the quantum states that are non-separable, the tests for and measures of entanglement are *consequential* on this definition.

1.3 Particle versus Mode Entanglement

These two papers deal with *identical* particles - bosons or fermions. In the *second quantisation* approach used here the system is regarded as a set of *quantum fields*, each of which may be considered as a collection of single particle states or *modes*. We now take into account the situation where systems of *identical particles* are involved. This requires us to give special consideration to the requirement that quantum states in such cases must conform to the *symmetrisation principle*. What sub-systems are possible must take into account that entanglement requires the specification of sub-systems that are *distinguishable* from each other and on which *measurements* can be made. In addition, the sub-systems must be able to exist as *separate* systems which can be prepared in quantum states for that sub-system alone. These key requirements that the sub-systems must be distinguishable, susceptible to measurements and can exist in separate quantum states are necessary for the concept of entanglement to make

physical sense, and have important consequences for the choice of sub-systems when identical particles are involved. These three key logical requirement for sub-systems rule out considering *labelled identical particles* as sub-systems and lead to the conclusion that sub-systems must be *modes*. Thus both the system and sub-systems will be specified via the modes that are involved, so here the sub-systems in terms of which non-entangled (and hence entangled) states are defined are *modes* or *sets* of modes, not particles [27], [28], [29]), [13], [44], [45]. In this approach, *particles* are associated with the *occupancies* of the various modes, so that situations with differing numbers of particles will be treated as differing quantum *states* of the same system, not as different systems - as in the *first quantisation* approach. Note that the choice of modes is *not unique* - original sets of orthogonal one particle states (modes) may be replaced by other orthogonal sets. An example is given in Section **2** of accompanying paper II. Modes can often be categorised as *localised* modes, where the corresponding single particle wavefunction is confined to a restricted spatial region, or may be categorised as *delocalised* modes, where the opposite applies. Single particle harmonic oscillator states are an example of localised modes, momentum states are an example of delocalised modes. This distinction is significant when phenomena such as EPR violations and teleportation are considered.

However, even if the system consists entirely of *distinguishable* particles we can still regard the sub-systems as collections of modes. Each distinguishable particle is still associated with a set of single particle states or modes (momentum eigenstates, harmonic oscillator eigenstates, **etc.**) that can be occupied. More general states associated with a single particle may be quantum superposition states of those with a single particle occupancy of the modes. If the overall system consists of a number of distinguishable particles each of which is considered as a sub-system, then each such sub-system can equally be regarded as the set of modes associated with the particular distinguishable particle. Overall system states involving just one particle of each type would be simultaneous eigenstates of the number operators for each of the distinguishable particles, with an eigenvalue of unity corresponding to there being only one particle of each type. The second quantisation approach can still be used, but is somewhat superfluous when the modes for each particle are only occupied once.

Although *multi-mode* systems are also considered, in this paper we mainly focus on *two mode* systems of identical *bosonic* atoms, where the atoms at most occupy only two single particle states or modes. For bosonic atoms this situation applies in two mode interferometry, where if a single hyperfine component is involved the modes concerned may be two distinct spatial modes, such as in a double well magnetic or optical trap, or if two hyperfine components are involved in a single well trap each component has its own spatial mode. Large numbers of bosons may be involved since there is no restriction on the number of bosons that can occupy a bosonic mode. For fermionic atoms each hyperfine component again has its own spatial mode. However, if large numbers of *fermionic* atoms are involved then as the Pauli exclusion principle only allows each mode to accommodate one fermion, it follows that a large number of modes must considered and two mode systems would be restricted to at most

two fermions. Consideration of multi-mode entanglement for large numbers of fermions is *outside* the scope of the present paper (see [46] for a treatment of this), and unless otherwise indicated the focus will be on *bosonic modes*. The paper focuses on identical bosonic *atoms* - whether the paper also applies to *photons* is less clear and will be discussed below.

1.4 Symmetrization and Super-Selection Rules

The work presented here begins with the *fundamental issue* of how an entangled state should be *defined* in the context of systems involving *identical particles*. To reiterate - in the commonly used *mathematical approach* for defining entangled states, this requires *first* defining a general *non-entangled* state, all *other* states therefore being entangled. We adhere to the original definition of Werner [5] in which the separable states are those that can be prepared by *local operations* and *classical communication* (*LOCC*). This approach is adopted by other authors, see for example [47], [24], [48]. However, in other papers - see for example [49], [50] so-called *separable non-local* states are introduced in which LOCC is *not* required (see SubSection.3.4.3 for an example). By contrast (and consistent with Werner's approach), in the present paper it is contended that the *density operators* both for the *quantum states* of the overall system and those for the non-entangled (local) sub-systems in the context of *non-entangled states* must be compatible with certain principles and rules that have been found to be both necessary and sufficient for understanding *physical experiments* in non-relativistic many-body systems. In some other work (discussed below) this has not been the case. A key feature required of all quantum states for systems involving *identical particles*, entangled or not is that they satisfy the *symmetrization principle* [51]. This places restrictions both on the form of the overall density operator and also on what can be validly considered to be a sub-system. In particular this rules out *individual* identical particles being treated as sub-systems, as is done in some papers (see below). If the system consists entirely of *distinguishable particles* then the symmetrization principle is not relevant. In addition, *super-selection rules* (SSR) [52] only allow density operators which have *zero coherences* between states with *differing total numbers* of particles to represent valid *quantum states*, and this will be taken into account for *all* quantum states of the overall system, entangled or not. This is referred to as the *global particle number super-selection rule* In *non-entangled* or *separable* states the density operator is a sum over products of sub-system density operators, each product being weighted by its probability of ocurring (see below for details). For the non-entangled or separable quantum states, a so-called *local particle number super-selection rule* will *also* be applied to the density operators describing each of the *sub-systems*. These sub-system density operators must then have have zero coherences between states with differing numbers of *sub-system particles*. This additional restriction excludes density operators as defining non-entangled states when the sub-system density operators do not conform to the local particle number super-selection rule. Consequently, density operators where the local particle number SSR does

not apply would be regarded as entangled states. This viewpoint is discussed in papers by Bartlett et al [47], [53] as one of several approaches for defining entangled states. However, other authors such as [49], [50] state on the contrary that states where the sub-system density operators do *not* conform to the local particle number super-selection rule *are* still separable, others such as [54], [55] do so by implication - the latter papers applied to atomic as well as photon modes. So in these two papers we are advocating a *different definition* to some *other definitions* of entanglement in identical particle systems, the consequence being that the set of entangled states is now much *larger.* This is a *key idea* in this paper - not only should super-selection rules on particle numbers be applied to the the *overall* quantum state, entangled or not, but it *also* should be applied to the density operators that describe states of the modal *sub-systems* involved in the general definition of *non-entangled* states.

Note that for systems entirely consisting of N *distinguishable* particles the super-selection rules are still true, but are now *superfluous.* Each sub-system is the set of modes or *one particle states* of the *specific* distinguishable particle and the overall state is an N particle state in which the sub-systems only contain *one* particle. Consequently there are no sub-system or system coherences between states with differing particle number.

The detailed reasons for adopting the viewpoint that the entanglement criteria be compliant with the requirement of the local particle number super-selecrtion rule (SSR) for the sub-system are set out below. As will be seen, the local particle number super-selection rule restriction *firstly* depends on the *fundamental requirement* that for *all* composite systems - whether identical particles are involved or not - non-entangled states are only those that can be prepared via processes that involve only *LOCC.* The requirement that the sub-system density operators in identical particle cases satisfy the local particle number SSR is *consequential* on the sub-system states being possible sub-system quantum states. As mentioned before, the general definition of non-entangled states based on LOCC preparation processes was first suggested by Werner [5]. Apart from the papers by Bartlett et al [47], [53] we are not aware that this LOCC/SSR based criteria for non-entangled states has been invoked previously for identical particle systems, indeed the opposite approach has been proposed [49], [50]. However, the idea of considering whether sub-system states should satisfy the local particle number SSR has been presented in several papers - [49], [50], [47], [53], [56], [57], [58], mainly in the context of pure states for bosonic systems, though in these papers the focus is on issues other than the definition of entanglement - such as quantum communication protocols [49], multicopy distillation [47], mechanical work and accessible entanglement [56], [57] and Bell inequality violation [58]. The consequences for entanglement of applying this super-selection rule requirement to the sub-system density operators are quite *significant,* and in the accompanying paper II important *new entanglement tests* are determined. Not only can it immediately be established that spin squeezing *requires* entangled states, but though several of the other inequalities (accompanying paper II) that have been used as signatures of entanglement are still valid, *additional tests* can be obtained which only apply to entangled states that

are defined to conform to the symmetrisation principle and the super-selection rules.

It is worth emphasising that requiring the sub-system density operators satisfy the local particle number SSR means that there are less states than otherwise would be the case which are classed as non-entangled, and *more states* will be regarded as *entangled*. It is therefore not surprising that additional tests for entanglement will result. If *further restrictions* are placed on the sub-system density operator - such as requiring them to correspond to a fixed number of bosons again there will be more states regarded as entangled, and even more entanglement tests will apply. A particular example is given in SubSection **4.3** of paper II, where the sub-systems are restricted to one boson states.

The *symmetrisation* requirement for systems involving identical particles is well established since the work of Dirac. There are two types of justification for applying the *super-selection rules* for systems of identical particles (both massive and otherwise). The *first* approach for invoking the superselection rule to exclude quantum superposition states with differing numbers of identical particles is based on simple considerations and may be summarised as:

1. No way is known for creating such SSR non-compliant states.

2. No way is known for measuring the properties of such states.

3. Coherence and interference effects can be understood in terms of SSR compliant states.

The *second* approach is more sophisticated and involves linking the absence or presence of SSR to whether or not there is a suitable *reference frame* in terms of which the quantum state is described [59], [60], [61], [49], [50], [62], [63], [64], [53], [56], [57], [34]. This approach will be described in SubSection 3.2 and Appendix 10, the key idea being that SSR are a consequence of considering the description of a quantum state by a real observer (Charlie) whose phase reference frame has an *unknown phase difference* from that of a hypothetical observer ((Alice), both studying the same system. Alice is assumed to possess a phase reference frame such that her description of the quantum state of the system violates the SSR. Charlie, on the other hand is an actual observer with no such phase reference frame. Thus, whilst Alice's description of the system involves a quantum state may violate the SSR, the description of the same system by Charlie will involve a quantum state that is SSR compliant. In the main part of this paper the density operator $\widehat{\rho}$ used to describe the various quantum states will be that of the external observer (Charlie). Note that if well-defined phase references do *exist* and the relationship between them is *known*, then the SSR can be challenged (see SubSection 3.3 and Appendix 10), but this situation does not apply in the case of massive bosons (or fermions).

It should be noted that both of these justifications for applying the SSR are dependent on what is practical in terms of measurements in *non-relativistic* quantum physics. Here the situation is clearer for systems of massive particles such as atoms than for massless particles such as photons. Applying SSR for photons is discussed in SubSubSection 3.2.9.

However, to allow for quantum states that as far as we know cannot be made or measured, and for which there are no known physical effects that require

15

their presence is an unnecessary feature to add to the non-relativistic quantum physics of many body systems or to quantum optics. Considerations based on the general principle of simplicity (Occam's razor) would suggest not doing so until a clear physical justification for including them is found. The quantum state is intended to specify what is known about a quantum system and how it was prepared. It is used to determine the probabilities for possible measurements on the system. Clearly there is no point in including non-SSR compliant terms in the density operator for the quantum state. Such terms would neither allow for possible preparation processes, or contribute to measurement probabilities associated with physical efects. Furthermore, experiments can be carried out on each of the mode sub-systems considered as a *separate* system, and essentially the *same reasons* that justify applying the super-selection rule to the overall system also apply to the separate mode sub-systems in the context of defining *non-entangled states*. Hence, unless it can be justified to ignore the super-selection rule for the overall system it would be *inconsistent* not to apply it to the sub-system as well. As we will see, for separable states the requirement that the overall state is SSR compliant generally implies that the sub-system states are SSR compliant - though in some special cases this is not the case (see SubSection 4.3.3 of paper II). The *onus* is on those who wish to ignore the super-selection rule for the separate sub-systems to justify why it is being applied to the overall system. In addition, *joint measurements* on *all* the sub-systems can be carried out, and the interpretation of the measurement probabilities requires the density operators for the sub-system states to be physically based. The general application of super-selection rules has however been challenged (see SubSection 3.2) on the basis that super-selection rules are not a fundamental requirement of quantum theory, but are restrictions that could be lifted if there is a suitable system that acts as a *reference* for the coherences involved. In Section 3 and in Appendix 11 an analysis of these objections to the super-selection rule is presented, and in Appendix 10 we see that the approach based on phase reference frames does indeed justify the application of the SSR both to the general quantum states for multi-mode systems of identical particles and to the sub-system states for non-entangled states of these systems.

The sceptic who wishes to ignore the super-selection rules in the definition of entanglement - and consequently only consider as valid tests for entanglement where SSR compliance is *not* used in their derivation - needs to carry out a research program analogous to that which resulted in *parity non-conservation* becoming a basic feature of *weak interaction theory*. The successful incorporation of parity non-conservation involved first proposing (on symmetry grounds) possible interactions in which parity was not conserved, second working out possible experiments that could confirm parity non conservation and third carrying out key experiments that did confirm this. At this stage no such work in regard to SSR violation in non-relativistic many body physics has been carried out or is likely to be in the near future (except possibly for photons). As we will see in paper II, none of the experimental methods for entanglement tests that we examine can detect SSR non-compliance - none involve a suitable *phase reference*. To ignore SSR in non-relativistic entanglement theory and experiment on

the grounds of scepticism would be analogous to ignoring parity conservation in quantum chemistry or atomic physics - areas which are well-understood in terms of parity being conserved (apart from the well-known parity violating effects of external electric fields). When and if SSR violation in non-relativistic many body physics is found would then be the time to revise the definition of quantum entanglement. In these two papers we will utilise the definition of entanglement and derive tests based on SSR compliance, though of course recognising that there are also tests that do not require SSR compliance which are *also* valid for SSR compliant states. Although other definitions of entanglement will be considered for comparison, to avoid confusion the SSR compliant definition will be the one which we mean when we refer to entanglement.

A further sound scientific argument can be presented in favour of studying SSR compliant entanglement tests (as is our aim in these papers (I) and (II)). This involves a consideration of what can be concluded from such tests by suppporters or sceptics of SSR. For example, one such test (see paper II) involves spin squeezing in two mode systems. If the state is separable *and* the sub-system states comply with local SSR then there is *no* spin squeezing. However, if experimental tests *do* demonstrate spin squeezing, then what can we conclude? The supporters of SSR compliance being required for the sub-systems would conclude that the state was not separable and hence *entanglement* is present between the subsystems. On the other hand, the sceptic who does not believe local SSR compliance is required would have no option but to conclude *either* that entanglement is present between the subsystems *or* (if they argue there is no entanglement) the state is separable but one or both of the quantum subsystems violates the SSR. The sceptic may favour the second conclusion, but that would then imply an *actual* experimental circumstance where superselection rules did *not* apply to the sub-system states. In that case, the issues raised in the last four paragraphs regarding *lack* of *phase references* or *SSR violating preparation processes* etc. must be addressed directly. Either way, the study of such SSR based experiments is clearly important. Put another way, suppose the sceptic were to derive a *different* test using the separability requirement *alone,* for which an experimental outcome shows that the two subsystems were indeed *not* entangled. This would seem to require a test for entanglement which is *necessary* as well as being sufficient - the latter alone being usually the case for entanglement tests. Such criteria and measurements are a challenge, but not impossible even though we have not met this challenge in these two papers. If the conclusion from the earlier *SSR based* experiment was *either* entanglement *or* separability with non SSR compliance, then if the result from the different test based *only* on separability ruled entanglement *out*, it follows that the system must be in a separable state in which the sub-system states violate the SSR. Conversely, the latter test may confirm the entanglement possibility found in the earlier test. Thus, *in principle* there could be a *pair* of experiments that give evidence of entanglement, *or* failure of the Super Selection Rule. For such investigations to be possible, the use of entanglement criteria that *do* invoke the local super-selection rules is *also* required.

1.5 Entanglement Tests and Experiments - Paper II

The main focus of the accompanying paper II is to derive the SSR compliant criteria and to consider the experimental implementation. This leads to important links between *spin squeezing* and entanglement. The link with *quantum correlation functions* (as proposed in Refs.[54], [55]) is also treated. Heisenberg Uncertainty Principle inequalities involving spin operators [65] and the consequent property of spin squeezing have been well-known in quantum optics for many years. The importance of spin squeezing in quantum metrology is discussed in the paper by Kitagawa et al [66] for general spin systems. It was suggested in this paper that correlations between the individual spins was needed to produce spin squeezing, though no quantitative proof was presented and the more precise concept of entanglement was not mentioned. For the case of two mode systems the earliest paper linking spin squeezing to entanglement is that of Sorensen et al [67], which considers a system of identical bosonic atoms, each of which can occupy one of two internal states. This paper states that spin squeezing requires the quantum state to be entangled, with a proof given in the Appendix. A consideration of how such spin squeezing may be generated via collisional interactions is also presented. The paper by Sorensen et al is often referred to as establishing the link between spin squeezing and entanglement - see for example Micheli et al [68], Toth et al [69], Hyllus et al [70]. However, the paper by Sorensen et al [67] is based on a definition of non-entangled states in which the sub-systems are the identical particles, and this is inconsistent with the symmetrization principle. However, the accompanying paper II establishes the link between spin squeezing and entanglement based on a definition of entanglement consistent with the system and sub-system density operators representing quantum states.

It is also important to consider which *components* of the spin operator vector are squeezed, and this issue is also considered in the accompanying paper. In the context of the present second quantisation approach to identical particle systems the three spin operator components for two mode systems are expressed in terms of the annihilation, creation operators for the two chosen modes. Spin squeezing can be defined (see Section **2** in the accompanying paper) in terms of the variances of these spin operators, however the *covariance matrix* for the three spin operators will in general have off-diagonal elements, and spin squeezing is also defined in terms of rotated spin operators referred to as *principal spin operators* for which the covariance matrix is *diagonal*. The principal spin operators are related to new mode annihilation, creation operators in the same form as for the original spin operators, where the *new modes* are two orthogonal linear combinations of the originally chosen modes. In discussing the relationship between spin squeezing and entanglement, the modes which may be entangled are generally those associated with the definition of the spin operators.

A further focus of the accompanying paper is on the relationship between entanglement and certain *correlation properties* of sub-system operators. Tests for entanglement based on such correlations have also been published - see for example [54], [55]. These tests were based on ignoring the super-selection rules,

18

so in the accompanying paper we present revised correlation tests for entanglement when the super-selection rules are definitely complied with. We also show the link between correlation tests and tests involving spin operators.

The accompanying paper also deals with the important question of what *measurement systems* are suitable for making spin and correlation tests for entanglement. We first consider a *simple two mode interferometer* which involves coupling the two modes employing a *resonant classical field pulse* which is associated with a variable *pulse area* for its amplitude and has an adjustable *phase*. It is shown that measurement of the mean value and variance of the *population difference* between the two modes *after* the interferometer pulse enables measurements of the *mean value* and *covariance matrix* elements of the *spin operators* for the quantum state that existed *before* the pulse was applied. The mean values and variances of certain spin operators are relevant for correlation and spin squeezing entanglement tests.

Paper II is focused on *two mode* bosonic systems. These are of particular interest because cold atomic gases cooled well below the Bose-Einstein condensation (BEC) transition temperature can be prepared where essentially only two modes are occupied ([71], [32]). This can be achieved for cases involving a single hyperfine components using a double well trap potential or for two hyperfine components using a single well. At higher temperatures more than two modes may be occupied, so *multi-mode* systems are also of importance and thus are considered in paper II.

1.6   Outlines of Papers I and II

The plan of the present paper is as follows. In Section 2 the key definitions of entangled states are covered, and a detailed discussion on why the symmetrisation principle and the super-selection rule is invoked in discussed in Section 3. Challenges to the necessity of the super-selection rule are outlined, with arguements against such challenges dealt with in Appendices 10 and 11. Two key mathematical inequalities are derived in Appendix 7. Details for the spin EPR paradox are given in Appendix 8. The final Section 4 summarises and discusses the key features about entanglement treated in this paper.

In the accompanying paper II, Section **2** sets out the definitions of spin squeezing and in the following Section **3** it is shown that spin squeezing is a signature of entanglement, both for the original spin operators with entanglement of the original modes and the principle spin operators with entanglement of the two new modes, and also for multi-mode cases. Details are in Appendices **A** and **B.** A number of other tests for entanglement proposed by other authors are considered in Sections **4, 5** and **6**, with details of these treatments set out in Appendices **B, C** and **D**. In Section **7** it is shown that a simple two mode interferometer can be used to measure the mean values and covariance matrix for the spin operators involved in entanglement tests. The treatment is then generalised to situations involving measurements on multi-mode systems. Details are covered in Appendices **G** and **H**.   Actual experiments aimed at detecting entanglement via spin squeezing tests are examined in Section **8.** The

final Section **9** summarises and discusses the key results regarding entanglement tests. Appendices **E, F** and **I** provide details regarding certain important states whose features are discussed in the paper - the "separable but non-local " states and the relative phase eigenstate.

2 Entanglement - General Features

2.1 Quantum States

The standard Copenhagen quantum theory notions of *physical systems* that can exist in various *states* and have associated *properties* on which *measurements* can be made are presumed in this paper. The measuring system may be also treated via quantum theory, but there is always some component that behaves *classically,* so that quantum fluctuations in the quantity recorded by the *observer* are small. The term *quantum state* (or "physical quantum state" or just "state" for short) refers to a state that can either be prepared via a process consistent with the laws of quantum physics and on which measurements can be then performed and the probabilistic results predicted from this state (*prediction*), or a state whose existence can be inferred from later quantum measurements (*retrodiction*). We may also refer to such states as *allowed* quantum states, and our approach is intended to be *physically* based. In quantum theory, quantum states are *represented* mathematically by *density operators* for mixed states or *state vectors* for pure states. For identical particle systems these representations must satisfy symmetrisation and other basic requirements in accordance with the laws of quantum theory. The probabilities of measurement outcomes and the probabilities associated with retrodiction can be interpreted as *Bayesian probabilities* [72], [73], and the quantum state is *observer dependent.* The quantum state, the system it is associated with and the quantities that can be measured are considered here as entities that are viewed as being *both* ontological and epistimological. Different observers may have different information about how the quantum state was prepared, hence the quantum state is in part *epistimological*, and would be described differently by different observers. Hence the observer is important, but as there is actually something out there to be studied, quantum states also have an *ontological* aspect. We will avoid the unqualified term "physical state" because this term is generally invoked in discussions about the pre-Copenhagen notion of *reality* and refers to some as yet unknown but more fundamental description of the system which underlies the quantum state [74]. Hidden variable theories attempt to describe this more fundamental physical state that is assumed to exist - attempts that so far have been unsuccessful if locality is also invoked (see below). In addition to those associated with physical quantum states, other density operators and state vectors may be introduced for mathematical convenience. For physical quantum states, the density operator is determined from either the preparation process or inferred from the measurement process **-** *quantum tomography* **-** and in general it is a statistical mixture of density operators for possible preparation processes. Measurement itself constitutes a possible preparation process. Following preparation, further experimental processes may change the quantum state and dynamical equations give the time evolution of the density operator between preparation and measurement, the simplest situation being where measurement takes place immediately after preparation. A full discussion of the predictive and retrodictive aspects of the density operator is given in papers by

Pegg et al [72], [75]. Whilst there are often different mathematical forms for the density operator that lead to the same predictive results for subsequent measurements, the results of the measurements can also be used to retrodictively determine the *preferred form* of the density operator that is consistent with the available preparation and measurement operators. An example is given in [75].

2.2 Entangled and Non-Entangled States

2.2.1 General Considerations

Here the commonly applied *physically-based approach* to mathematically defining entangled states will be described [14]. The definition involves vectors and density operators that represent states than can be prepared in real experiments, so the mathematical approach is to be *physically based*. The concept of quantum entanglement involves *composite systems* made up of component *sub-systems* each of which are *distinguishable* from the other sub-systems, and where each could constitute a stand-alone quantum system. This means the each sub-system will have its own set of physically realisable quantum states - mixed or pure - which could be prepared independently of the quantum states of the other sub-systems. As will be seen, the requirement that sub-systems be distinguishable and their states be physically preparable will have important consequences, especially in the context of identical particle systems. The formal definition of what is meant by an entangled state starts with the pure states, described via a vector in a Hilbert space. The formalism of quantum theory allows for *pure states* for composite systems made up of two or more distinct sub-systems via tensor products of sub-system states

$$|\Phi\rangle = |\Phi_A\rangle \otimes |\Phi_B\rangle \otimes |\Phi_C\rangle \, ... \tag{1}$$

Such products are called *non-entangled* or *separable* states. However, since these product states exist in a Hilbert space, it follows that linear combinations of such products of the form

$$|\Phi\rangle = \sum_{\alpha\beta\gamma..} C_{\alpha\beta\gamma..} |\Phi_A^\alpha\rangle \otimes \left|\Phi_B^\beta\right\rangle \otimes |\Phi_C^\gamma\rangle . \tag{2}$$

could also represent possible pure quantum states for the system. Such *quantum superpositions* which cannot be expressed as a *single* product of sub-system states are known as *entangled* (or *non-separable*) states.

The concept of entanglement can be extended to *mixed states*, which are described via density operators in the Hilbert space. If A, B, ... are the sub-systems with $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, being density operators the sub-systems A, B, .then a *general non-entangled* or *separable* state is one where the overall density operator $\widehat{\rho}$ can be written as the weighted sum of tensor products of these sub-system density operators in the form [5]

$$\widehat{\rho} = \sum_R P_R \, \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes \widehat{\rho}_R^C \otimes ... \tag{3}$$

22

with $\sum_R P_R = 1$ and $P_R \geq 0$ giving the probability that the specific product state $\widehat{\rho}_R = \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes \widehat{\rho}_R^C \otimes ...$ occurs. It is assumed that at least in principle such separable states can be prepared [5]. This implies the possibility of turning off the interactions between the different sub-systems, a task that may be difficult in practice except for well-separated sub-systems. *Entangled* states (or *non-separable* states) are those that cannot be written in this form, so in this approach knowing what the term entangled state refers to is based on *first* knowing what the general form is for a non-entangled state. The density operator $\widehat{\rho} = |\Phi\rangle \langle\Phi|$ for the pure state in (2) is not of the form (3), as there are cross terms of the form $C_{\alpha\beta\gamma..} C_{\theta\lambda\eta..}^* (|\Phi_A^\alpha\rangle \langle\Phi_A^\theta|) \otimes (|\Phi_B^\beta\rangle \langle\Phi_B^\lambda|) \otimes ...$ involved.

The concepts of separability and entanglement based on the Eqs. (1) and (3) for non-entangled states do not however just rest on the mathematical forms alone. Implicitly there is the *assumption* that separable quantum states described by the two expressions can actually be created in *physical processes*. The sub-systems involved must therefore be *distinguishable quantum systems* in their own right, and the sub-system states $|\Phi_A\rangle, |\Phi_B\rangle, ...$ or $\widehat{\rho}_R^A, \widehat{\rho}_R^B, ...$ must also be *possible quantum states* for the sub-systems. We will return to these requirements later. The issue of the physical preparation of non-entangled (separable) states starting from some uncorrelated fiducial state for the separate sub-systems was introduced by Werner [5], and discussed further by Bartlett et al (see [47], Section IIB). This involves the ideas of *local operations* and *classical communication (LOCC)* dealt with in the next SubSection.

The key requirement is that entangled states exhibit a novel *quantum feature* that is only found in *composite* systems. Separable states are such that the *joint probability* for measurements of all physical quantities associated with the sub-systems can be found from separate measurement probabilities obtained from the sub-system density operators $\widehat{\rho}_R^A, \widehat{\rho}_R^B$, etc and the overall classical probability P_R (see SubSection2.3). This feature of separable probabilities is absent in *certain* entangled states, and because of this key *non-separability feature* Schrodinger called these states "entangled". The separability feature for the joint probabilities is essentially a classical feature and applies in hidden variable theories (HVT) (see SubSection2.5) applied to quantum systems - as well as to quantum separable states. The fact that entangled states are quantum states that can exhibit the failure of this separability feature for classical LHV theories highlights entanglement being a non-classical feature for composite systems.

An alternative *operational approach* to defining entangled states focuses on whether or not they exhibit certain non-classical features such as Bell Inequality violation or whether they satisfy certain mathematical tests such as having a non-negative partial transpose [41], [30], and a *utilititarian approach* focuses or whether entangled states have technological applications such as in various quantum information protocols. As will be seen in SubSection 3.4, the particular definition of entangled states based on their non-creatability via LOCC essentially coincides with the approach used in the present paper. It has been realised for some time that different types of entangled states occur, for example states in which a Bell inequality is violated or states demonstrating an EPR paradox

[76]. Wiseman et al [23], [24], [25] and Reid et al [77], [15], [16], [26] discuss the concept of a *heirarchy* of *entangled states*, with states exhibiting *Bell nonlocality* being a subset of states for which there is *EPR steering*, which in turn is a subset of all the *entangled states*, the latter being defined as states whose density operators cannot be written as in Eq. (3) though without further consideration if additional properties are required for the sub-system density operators. The operational approach could lead into a quagmire of differing interpretations of entanglement dependng on which non-classical feature is highlighted, and the utilitarian approach implies that all entangled states have a technological use - which is by no means the case. For these reasons, the present physical approach based on the quantities involved representing allowed sub-system states is generally favoured [14]. It is also compatible with later classifying entangled states in a heirarchy.

2.2.2   Local Systems and Operations

As pointed out by Vedral [13], one reason for calling states such as in Eqs.(1) and (3) separable is associated with the idea of performing operations on the separate sub-systems that do not affect the other sub-systems. Such operations on such *local systems* are referred to as *local operations* and include unitary operations $\widehat{U}_A$, $\widehat{U}_B$, that change the states via $\widehat{\rho}_R^A \rightarrow \widehat{U}_A \widehat{\rho}_R^A \widehat{U}_A^{-1}$, $\widehat{\rho}_R^B \rightarrow \widehat{U}_B \widehat{\rho}_R^B \widehat{U}_B^{-1}$, etc as in a time evolution, and could include processes by which the states $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, are separately prepared from suitable initial states.

We note that performing local operations on a separable state only produces another separable state, not an entangled state. Such local operations are obviously faciltated in experiments if the sub-systems are essentially *non-interacting* - such as when they are spatially *well-separated*, though this does not have to be the case. The local systems and operations could involve sub-systems whose quantum states and operators are just in different parts of Hilbert space, such as for cold atoms in different hyperfine states even when located in the same spatial region. Note the distinction between *local* and *localised*. As described by Werner [5], if one observer (Alice) is associated with preparing separate sub-system A in an allowed quantum state $\widehat{\rho}_R^A$ via local operations with a probability P_R, a second observer (Bob) could be then advised via a *classical communication* channel to prepare sub-system B in state $\widehat{\rho}_R^B$ via local operations. After repeating this process for different choices R of the correlated pairs of sub-system states, the overall quantum state prepared by both observers via this local operation and classical communication protocol *(LOCC)* would then be the bipartite non-entangled state $\widehat{\rho} = \sum_R P_R \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B$. Multipartite non-entangled states of the form (3) can also be prepared via LOCC protocols involving further observers. As will be seen, the separable or non-entangled states are just those that can be prepared by LOCC protocols.

2.2.3 Constraints on Sub-System Density Operators

A key issue however is whether density operators $\widehat{\rho}$ and $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, in Eq. (3) always represent possible *quantum states*, even if the operators $\widehat{\rho}$ and $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, etc satisfy all the standard mathematical requirements for density operators - Hermitiancy, positiveness, trace equal to unity, trace of density operator squared being not greater than unity. In this paper it will be argued that for systems of identical massive particles there are further requirements not only on the overall density operator, but also (for separable states) on those for the individual sub-systems that are imposed by *symmetrisation* and *super-selection* rules.

2.2.4 Classical Entanglement

In addition to quantum entanglement there is a body of work (see [78], [79]), [80] dealing with so-called *classical entanglement*. Here the *states* of *classical* systems - such as a classical EM field - are represented via a formalism involving *linear vector spaces* and classical entanglement is defined *mathematically*. A discussion of classical entanglement is beyond the scope of this paper. Although there are some formal similarities with quantum entanglement - and even Bell type inequalites which can be violated, there are key features that is not analogous to that for composite quantum systems - quantum non-locality being one [78]. In the end, rather than just focusing on similarities in the mathematical formalisms, classical and quantum entanglement are seen as fundamentally different when the physics of the two different types of system - one classical and deterministic, the other quantum and probabilistic are taken into account. In particular, the key feature of quantum entanglement relating to joint measurement probabilities is quite different to the corresponding one for classical entanglement.

2.3 Separate and Joint Measurements, Reduced Density Operator

In this SubSection we consider separate and joint measurements on systems involving several sub-systems and introduce results for probabilities, mean values for measurements on one of the sub-systems which are conditional on the results for measurements on another of the sub-systems. This will require consideration of quantum theoretical *conditional probabilities*. The measurements involved will be assumed for simplicity to be von Neumann *projective measurements* for physical quantities represented by Hermitian operators $\widehat{\Omega}$, which project the quantum state into subspaces for the eigenvalue λ_i that is measured, the subspaces being associated with Hermitian, idempotent *projectors* $\widehat{\Pi}_i$ whose sum over all eigenvalues is unity. These concepts are treated in several quantum theory textbooks, for example [11], [81]. For completeness, an account setting out the key results is presented in Appendix 6.

2.3.1   Joint Measurements on Sub-Systems

For situations involving distinct sub-systems measurements can be carried out on all the sub-systems and the results expressed in terms of the *joint probability* for various outcomes. If $\widehat{\Omega}_A$ is a physical quantity associated with sub-system A, with eigenvalues λ_i^A and with $\widehat{\Pi}_i^A$ the projector onto the subspace with eigenvalue λ_i^A, $\widehat{\Omega}_B$ is a physical quantity associated with sub-system B, with eigenvalues λ_j^B and with $\widehat{\Pi}_j^B$ the projector onto the subspace with eigenvalue λ_j^B etc., then the *joint probability* $P_{AB..}(i, j, ...)$ that measurement of $\widehat{\Omega}_A$ leads to result λ_i^A , measurement of $\widehat{\Omega}_B$ leads to result λ_j^B ,etc is given by

$$P_{AB..}(i, j, ...) = Tr(\widehat{\Pi}_i^A \, \widehat{\Pi}_j^B \, ...\widehat{\rho}) \tag{4}$$

This joint probability depends on the full density operator $\widehat{\rho}$ representing the allowed quantum state as well as on the quantities being measured. Here the projectors (strictly $\widehat{\Pi}_i^A \otimes \widehat{1}^B \otimes ...$, $\widehat{1}^A \otimes \widehat{\Pi}_j^B \otimes ...$, etc) commute, so the order of measurements is immaterial. An alternative notation in which the physical quantities are also specified is $P_{AB..}(\widehat{\Omega}_A, i; \widehat{\Omega}_B, j; ...)$.

2.3.2   Single Measurements on Sub-Systems and Reduced Density Operator

The *reduced density operator* $\widehat{\rho}_A$ for sub-system A given by

$$\widehat{\rho}_A = Tr_{B,C,...}(\widehat{\rho}) \tag{5}$$

and enables the results for measurements on sub-system A to be determined for the situation where the results for all joint measurements involving the other sub-systems are *discarded*. The probability $P_A(i)$ that measurement of $\widehat{\Omega}_A$ leads to result λ_i^A irrespective of the results for meaurements on the other sub-systems is given by

$$
\begin{aligned}
P_A(i) &= \sum_{j,k,...} P_{AB..}(i, j, ...) \\
&= Tr(\widehat{\Pi}_i^A \, \widehat{\rho}) \tag{6} \\
&= Tr_A(\widehat{\Pi}_i^A \, \widehat{\rho}_A) \tag{7}
\end{aligned}
$$

using $\sum_j \widehat{\Pi}_j^B = \widehat{1}$, etc. Hence the reduced density operator $\widehat{\rho}_A$ plays the role of specifying the quantum state for mode A considered as a separate sub-system, even if the original state $\widehat{\rho}$ is entangled. An alternative notation in which the physical quantity is also specified is $P_A(\widehat{\Omega}_A, i)$.

2.3.3 Mean Value and Variance

The *mean value* for measuring a physical quantity $\widehat{\Omega}_A$ will be given by

$$
\begin{aligned}
\left\langle \widehat{\Omega}_A \right\rangle &= \sum_{\lambda_i^A} \lambda_i^A P_A(i) \\
&= Tr_A(\widehat{\Omega}^A \widehat{\rho}_A)
\end{aligned}
\tag{8}
$$

where we have used $\widehat{\Omega}^A = \sum_{\lambda_i^A} \lambda_i^A \widehat{\Pi}_i^A$.

The *variance* of measurements of the physical quantity $\widehat{\Omega}_A$ will be given by

$$
\begin{aligned}
\left\langle (\Delta \widehat{\Omega}^A)^2 \right\rangle &= \sum_{\lambda_i^A} (\lambda_i^A - \left\langle \widehat{\Omega}_A \right\rangle)^2 P_A(i) \\
&= Tr_A\left(\left(\widehat{\Omega}^A - \left\langle \widehat{\Omega}_A \right\rangle \right)^2 \widehat{\rho}_A \right)
\end{aligned}
\tag{9}
$$

so both the mean and variance only depend on the reduced density operator $\widehat{\rho}_A$.

On the other hand the *mean value* of a *product* of sub-system operators $\widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes ...$, where $\widehat{\Omega}_A, \widehat{\Omega}_B, \widehat{\Omega}_C, ..$ are Hermitian operators representing physical quantities for the separate sub-systems, is given by

$$
\begin{aligned}
\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes . \right\rangle &= \sum_{\lambda_i^A} \sum_{\lambda_j^B} ... \lambda_i^A \lambda_j^B ... P_{AB..}(i, j, ...) \\
&= Tr\left(\widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes . \right) \widehat{\rho}
\end{aligned}
\tag{10}
$$

which involves the overall system density operator, as expected.

2.3.4 Conditional Probabilities

Treating the case of two sub-systems for simplicity we can use Bayes theorem (see Appendix 6, Eq.(155)) to obtain expressions for *conditional probabilities* [14]. The conditional probability that if measurement of $\widehat{\Omega}_B$ associated with sub-system B leads to eigenvalue λ_j^B then measurement of $\widehat{\Omega}_A$ associated with sub-system A leads to eigenvalue λ_i^A is given by

$$
P_{AB}(i|j) = Tr(\widehat{\Pi}_i^A \widehat{\Pi}_j^B \widehat{\rho})/Tr(\widehat{\Pi}_j^B \widehat{\rho})
\tag{11}
$$

In general, the overall density operator is required to determine the conditional probability. An alternative notation in which the physical quantities are also specified is $P_{AB}(\widehat{\Omega}_A, i|\widehat{\Omega}_B, j)$.

As shown in Appendix 6 the conditional probability is given by

$$
P_{AB}(i|j) = Tr(\widehat{\Pi}_i^A \widehat{\rho}_{cond}(\widehat{\Omega}_B, \lambda_j^B))
\tag{12}
$$

where

$$
\widehat{\rho}_{cond}(\widehat{\Omega}_B, \lambda_j^B) = \widehat{\Pi}_j^B \widehat{\rho} \widehat{\Pi}_j^B / Tr(\widehat{\Pi}_j^B \widehat{\rho})
\tag{13}
$$

is the so-called *conditioned density operator*, corresponding the quantum state produced following the measurement of $\widehat{\Omega}_B$ that obtained the result λ_j^B. The conditional probability result is the same as

$$P_{AB}(i|j) = Tr(\widehat{\Pi}_i^A \widehat{\rho}_{cond}(\widehat{\Omega}_B, \lambda_j^B)) \tag{14}$$

which is the same as the expression (6) with $\widehat{\rho}$ replaced by $\widehat{\rho}_{cond}(\widehat{\Omega}_B, \lambda_j^B)$. This is what would be expected for a conditioned measurement probability.

Also, if the measurement results for $\widehat{\Omega}_B$ are not recorded the conditioned density operator now becomes

$$
\begin{aligned}
\widehat{\rho}_{cond}(\widehat{\Omega}_B) &= \sum_{\lambda_j^B} P_B(j) \widehat{\rho}_{cond}(\widehat{\Omega}_B, \lambda_j^B) \\
&= \sum_{\lambda_j^B} \widehat{\Pi}_j^B \widehat{\rho} \widehat{\Pi}_j^B
\end{aligned} \tag{15}
$$

This is still different to the original density operator $\widehat{\rho}$ because a measurement of $\widehat{\Omega}_B$ has occured, even if we dont know the outcome. However, the measurement probability for $\widehat{\Omega}_A$ is now

$$
\begin{aligned}
P_{AB}(i|Any\,j) &= Tr(\widehat{\Pi}_i^A \widehat{\rho}_{cond}(\widehat{\Omega}_B)) \\
&= Tr(\widehat{\Pi}_i^A \widehat{\rho}) \tag{16} \\
&= P_A(i) \tag{17}
\end{aligned}
$$

where we have used the cyclic properties of the trace, $\left(\widehat{\Pi}_j^B\right)^2 = \widehat{\Pi}_j^B$ and $\sum_{\lambda_j^B} \widehat{\Pi}_j^B = \widehat{1}$. The results in Eqs. (16) and (17) are the same as the measurement probability for $\widehat{\Omega}_A$ if no measurement for $\widehat{\Omega}_B$ had taken place at all. This is perhaps not surprising, since the record of the latter measurements was discarded. Another way of showing this result is that Bayes Theorem tells us that $\sum_j P_{AB}(i|j) P_B(j) = \sum_j P_{AB}(i,j) = P_A(i)$, since $\sum_j P_{AB}(i,j)$ is the probability that measurement of $\widehat{\Omega}_A$ will lead to λ_i^A and measurement of $\widehat{\Omega}_B$ will lead to any of the λ_j^B. This result is called the no-signalling theorem [14].

Also, as $P_{AB}(i|Any\,j) = Tr(\widehat{\Pi}_i^A \widehat{\rho}_{cond}(\widehat{\Omega}_B))$ we see from (7) that

$$\widehat{\rho}_A = Tr_B(\widehat{\rho}_{cond}(\widehat{\Omega}_B)) \tag{18}$$

showing that the trace over B of the conditioned density operator for the state obtained by measuring *any* observable $\widehat{\Omega}_A$ and then discarding the results just gives the *reduced density operator* for sub-system A.

2.3.5   Conditional Mean and Variance

As explained in Appendix 6, to determine the *conditioned mean value* of $\widehat{\Lambda}$ after measurement of $\widehat{\Omega}$ has led to the eigenvalue λ_i we use $\widehat{\rho}_{cond}(\widehat{\Omega}, i)$ rather than

$\widehat{\rho}$ in the mean formula $\left\langle\widehat{\Lambda}\right\rangle = Tr(\widehat{\Lambda}\widehat{\rho})$ and the result is given in terms of the conditional probability $P(\widehat{\Lambda}j|\widehat{\Omega}i)$. Here we refer to two commuting observables and include the operators in the notation to avoid any misinterpretation. Hence

$$
\begin{aligned}
\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega},i} &= Tr(\widehat{\Lambda}\widehat{\rho}_{cond}(\widehat{\Omega},i)) \\
&= \sum_{j}\mu_{j}\,P(\widehat{\Lambda},j|\widehat{\Omega},i)
\end{aligned}
\tag{19}
$$

For the *conditioned variance* of $\widehat{\Lambda}$ after measurement of $\widehat{\Omega}$ has led to the eigenvalue λ_i we use $\widehat{\rho}_{cond}(\widehat{\Omega},i)$ rather than $\widehat{\rho}$ and the conditioned mean $\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega},i}$ rather than $\left\langle\widehat{\Lambda}\right\rangle$ in the variance formula $\left\langle\Delta\widehat{\Lambda}^2\right\rangle = Tr((\widehat{\Lambda}-\left\langle\widehat{\Lambda}\right\rangle)^2\widehat{\rho})$. Hence

$$
\begin{aligned}
\left\langle\Delta\widehat{\Lambda}^2\right\rangle_{\widehat{\Omega},i} &= Tr((\widehat{\Lambda}-\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega},i})^2\widehat{\rho}_{cond}(\widehat{\Omega},i)) \\
&= \sum_{j}(\mu_{j}-\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega},i})^2\,P(\widehat{\Lambda},j|\widehat{\Omega},i)
\end{aligned}
\tag{20}
$$

If we weighted the conditioned mean by the probability $P(\widehat{\Omega},i)$ that measuring $\widehat{\Omega}$ has led to the eigenvalue λ_i and summed over the possible outcomes λ_i for the $\widehat{\Omega}$ measurement, then we obtain the mean for measurements of $\widehat{\Lambda}$ after un-recorded measurements of $\widehat{\Omega}$ have occured. From Bayes theorem $\sum_{i}P(\widehat{\Lambda},j|\widehat{\Omega},i)P(\widehat{\Omega},i) = P(\widehat{\Lambda},j)$ so this gives the *unrecorded mean* $\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega}}$ as

$$
\begin{aligned}
\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega}} &= \sum_{i}\left\langle\widehat{\Lambda}\right\rangle_{\widehat{\Omega},i}P(\widehat{\Omega},i) \\
&= \sum_{j}\mu_{j}\,P(\widehat{\Lambda},j) \\
&= \left\langle\widehat{\Lambda}\right\rangle
\end{aligned}
\tag{21}
$$

which is the usual mean value for measurements of $\widehat{\Lambda}$ when no measurements of $\widehat{\Omega}$ have occured. Note that no such similar result occurs for the *unrecorded variance* $\left\langle\Delta\widehat{\Lambda}^2\right\rangle_{\widehat{\Omega}}$

$$
\begin{aligned}
\left\langle\Delta\widehat{\Lambda}^2\right\rangle_{\widehat{\Omega}} &= \sum_{i}\left\langle\Delta\widehat{\Lambda}^2\right\rangle_{\widehat{\Omega},i}P(\widehat{\Omega},i) \\
&\neq \left\langle\Delta\widehat{\Lambda}^2\right\rangle
\end{aligned}
\tag{22}
$$

2.4　Non-Entangled States

In this SubSection we will set out the key results for measurements on non-entangled states.

2.4.1　Non-Entangled States - Joint Measurements on Sub-Systems

In the case of the general *non-entangled state* we find that the joint probability is

$$P_{AB..}(i, j, ...) = \sum_R P_R \, P_A^R(i) P_B^R(j) \, ... \tag{23}$$

where

$$P_A^R(i) = Tr(\widehat{\Pi}_i^A \, \widehat{\rho}_R^A) \qquad P_B^R(j) = Tr(\widehat{\Pi}_j^B \, \widehat{\rho}_R^B) \qquad .. \tag{24}$$

are the probabilities for measurement results for $\widehat{\Omega}_A$, $\widehat{\Omega}_B$, ... on the separate sub-systems with density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, etc and the overall joint probability is given by the products of the probabilities $P_A^R(i)$, $P_B^R(j)$, .. for the measurement results λ_i^A, λ_j^B, ... for physical quantities $\widehat{\Omega}_A$, $\widehat{\Omega}_B$, ... if the sub-systems are in the states $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, etc. Note that here $P_A^R(i)$, $P_B^R(j)$, are given by quantum theory formulae for the sub-systenm states. For simplicity only quantized measured values will be considered - the extension to continuous values is straightforward. Thus the results for the probabilities of joint measurements when the system is in a separable quantum state are determined by the measurement probabilities in *possible* quantum states for the sub-systems, combined with a *classical* probability for creating the particular pair of sub-system quantum states. Note the emphasis on "possible" - some of the separable states described in [49] are not possible.

Furthermore, if we consider measurements of the physical quantity $\widehat{\Omega}_A \otimes \widehat{\Omega}_B$ then for a separable state the *mean value* for measurement of this quantity is given by

$$\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \right\rangle = Tr\left(\widehat{\Omega}_A \otimes \widehat{\Omega}_B \, \widehat{\rho} \right) = \sum_R P_R \left\langle \widehat{\Omega}_A \right\rangle_R^A \left\langle \widehat{\Omega}_B \right\rangle_R^B \tag{25}$$

where $\left\langle \widehat{\Omega}_A \right\rangle_R^A = Tr_a\left(\widehat{\Omega}_A \, \widehat{\rho}_R^A \right)$ and $\left\langle \widehat{\Omega}_B \right\rangle_R^B = Tr_b\left(\widehat{\Omega}_B \, \widehat{\rho}_R^B \right)$ are the mean values of $\widehat{\Omega}_A$ and $\widehat{\Omega}_B$ for the sub-system states $\widehat{\rho}_R^A$ and $\widehat{\rho}_R^B$ respectively. If $\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \right\rangle = \left\langle \widehat{\Omega}_A \right\rangle \left\langle \widehat{\Omega}_B \right\rangle$ then the state is said to be *uncorrelated*. Separable states are *correlated* except for the case where $\widehat{\rho}_{sep} = \widehat{\rho}^A \otimes \widehat{\rho}^B$, but the correlation is essentially *non-quantum* and attributable to the classical probabilities P_R. However, for separable states the inequality $|\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \right\rangle|^2 \leq \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle$ applies, so that if $|\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \right\rangle|^2 > \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \right\rangle$ then the state is entangled.

In the simple non-entangled *pure state* situation in Eq.(1) the joint probabilty only involves a single product of sub-system probabilities

$$P_{AB..}(i, j, ...) = P_A(i) P_B(j)... \tag{26}$$

where

$$P_A(i) = \langle \Phi_A | \, \widehat{\Pi}_i^A \, | \Phi_A \rangle \qquad P_B(j) = \langle \Phi_B | \, \widehat{\Pi}_j^B \, | \Phi_B \rangle \qquad .. \qquad (27)$$

just give the probabilities for measurements in the separate sub-systems.

This *key result* (23) showing that the joint measurement probability for a separable state only depends on *separate measurement probabilities* for the sub-systems, together with the classical probability for preparing correlated product states of the sub-systems, does *not* necessarily apply for entangled states [5]. However the *key quantum feature* for *composite systems* of non-separability for joint measurement probabilites applies only to entangled states. This strange quantum feature of entangled states has been regarded as particularly unusual when the sub-systems are *spatially well-separated* (or non-local) because then measurement events can become space-like separated. This is relevant to quantum paradoxes such as Einstein-Poldolsky-Rosen (EPR) and Bell's theorem which aim to show there could be no causal classical theory explaining quantum mechanics [1], [2]. Measurements on sub-system A of physical quantity $\widehat{\Omega}_A$ affect the results of measurements of $\widehat{\Omega}_B$ at the same time on a distant sub-system B, even if the choice of measured quantity $\widehat{\Omega}_B$ is unknown to the experimenter measuring $\widehat{\Omega}_A$. As will be shown below, a similar result to (23) also occurs in *hidden variable theory* - a classical theory - so non-separability for joint measurements resulting from entanglement is a truly *non-classical feature* of composite systems.

2.4.2   Non-Entangled States - Single Sub-System Measurements

For the general non-entangled state, the reduced density operator for sub-system A is given by

$$\widehat{\rho}_A = \sum_R P_R \, \widehat{\rho}_R^A \qquad (28)$$

A key feature of a non-entangled state is that the results of a measurement on any *one* of the sub-systems is *independent* of the states for the *other* subsystems. From Eqs.(7) and (28) the probability $P_A(i)$ that measurement of $\widehat{\Omega}_A$ leads to result λ_i^A is given by

$$P_A(i) = \sum_R P_R \, P_A^R(i) \qquad (29)$$

where the reduced density operator is given by Eq. (28) for the non-entangled state in Eq. (3). This result only depends on the reduced density operator $\widehat{\rho}_A$, which represents a state for sub-system A and which is a statistical mixture of the sub-system states $\widehat{\rho}_R^A$, with a probability P_R that is the *same* for all sub-systems. The result for the measurement probability $P_A(i)$ is just the statistical average of the results that would apply if sub-system A were in possible states $\widehat{\rho}_R^A$. For all quantum states the final expression for the measurement probability $P_A(i)$ only involves a trace of quantities $\widehat{\Pi}_i^A$, $\widehat{\rho}_A$ that apply to sub-system A, but for a non-entangled state the reduced density operator $\widehat{\rho}_A$ is given by an expression (28) that does *not* involve density operators for the other subsystems. Thus for a non-entangled state, the probability $P_A(i)$ is *independent*

31

of the states $\widehat{\rho}_R^B$, $\widehat{\rho}_R^C$, associated with the other sub-systems. Analogous results apply for measurements on the other sub-systems.

2.4.3   Non-Entangled States - Conditional Probability

For a general non-entangled bipartite mixed state the conditional probability is given by

$$P_{AB}(i|j) = \sum_R P_R\, P_A^R(i) P_B^R(j) / \sum_R P_R\, P_B^R(j) \tag{30}$$

which in general depends on $\widehat{\Omega}_B$ associated with sub-system B and the eigenvalue λ_j^B. This may seem surprising for the case where A and B are localised sub-systems which are well separated. Even for separable states a measurement result for sub-system B will give *immediate* information about a totally separated measurement on sub-system A - which is *space-like* separated. However it should be remembered that the general separable system can still be a *correlated* state, since each sub-system density operator $\widehat{\rho}_R^B$ for sub-system B is matched with a corresponding density operator $\widehat{\rho}_R^A$ for sub-system A. Results at A can be correlated with those at B, so the observer at A can potentially infer from a local measurement on the sub-system A the result of a local measurement on sub-system B. It is therefore not necessarily the case that measurement results for A are independent of those for B. However, as we will see below, such correlations (usually) hve a classical interpretation. Result (30) is *not* a case of the "spooky action at a distance" that Einstein [1] referred to.

However, for a non-entangled pure state where $\widehat{\rho} = \widehat{\rho}^A \otimes \widehat{\rho}^B$ we do find that

$$P_{AB}(i|j) = P_A(i) \tag{31}$$

where $P_A(i) = Tr(\widehat{\Pi}_i^A \widehat{\rho}^A)$. For separable pure states the conditional probability is independent of $\widehat{\Omega}_B$ associated with sub-system B and the eigenvalue λ_j^B.

Also of course $\sum_j P_{AB}(i|j) P_B(j) = P_A(i)$ is true for separable states since it applies to general bipartite states. Hence if the measurement results for $\widehat{\Omega}_B$ are discarded then the probability distribution for measurements on $\widehat{\Omega}_A$ will be determined from the conditioned density operator $\widehat{\rho}_{cond}(\widehat{\Omega}_B)$ and just result in $P_A(i)$ - as in shown in Eq.(17) for any quantum state.

2.4.4   Non-Entangled States - Mean Values and Correlations

For non-entangled states as in Eq. (3) the mean value for measuring a physical quantity $\widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes ...$, where $\widehat{\Omega}_A$, $\widehat{\Omega}_B$, $\widehat{\Omega}_C$, .. are Hermitian operators representing physical quantities for the separate sub-systems can be obtained from Eqs.(3) and (10) and is given by

$$\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes . \right\rangle = \sum_R P_R \left\langle \widehat{\Omega}_A \right\rangle_R^A \left\langle \widehat{\Omega}_B \right\rangle_R^B \left\langle \widehat{\Omega}_C \right\rangle_R^C ... \tag{32}$$

where

$$\left\langle \widehat{\Omega}_K \right\rangle_R^K = Tr(\widehat{\Omega}_K \, \widehat{\rho}_R^K), \qquad (K = A, B, ...) \qquad (33)$$

is the mean value for measuring $\widehat{\Omega}_K$ in the K sub-system when its density operator is $\widehat{\rho}_R^K$. Since the overall mean value is not equal to the product of the separate mean values, the measurements on the sub-systems are said to be *correlated*. However, for the general non-entangled state as the mean value is just the products of mean values weighted by the probability of preparing the particular product state - which involves a LOCC protocal, as we have seen - the correlation is *classical* rather than *quantum* [14]. In the case of a single product state where $\widehat{\rho} = \widehat{\rho}^A \otimes \widehat{\rho}^B \otimes \widehat{\rho}^C \otimes ...$ we have $\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes . \right\rangle = \left\langle \widehat{\Omega}_A \right\rangle^A \left\langle \widehat{\Omega}_B \right\rangle^B \left\langle \widehat{\Omega}_C \right\rangle^C ...$ which is just the product of mean values for the separate sub-systems, and in this case the measurements on the sub-systems are said to be *uncorrelated*. For entangled states however the last result for $\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes . \right\rangle$ does not apply, and the correlation is strictly quantum.

2.5   Local Hidden Variable Theories

In a general *local hidden variable theory* as envisaged by Einstein et al [1] and Bell [3], physical quantities associated with the sub-systems are denoted Ω_A, Ω_B etc, which are real numbers not operators. Their values are assumed to be λ_i^A, λ_j^B etc - having the same ranges as in quantum theory, since HVT does not challenge the quantization feature. In the *realist* viewpoint of HVT all the physical quantities have *definite values* at *any* time, the probabilities for measuring these values being determined from a set of hidden variables ξ, which are themselves given by a probability function $P(\xi)$ for each state preparation process. Measurement is *not* required for the values for physical quantities to be created, as in quantum theory, nor do the hidden variables change as a result of the act of measurement itself (though they may change as a result of local interactions of the system with the measurement apparatus [82], [83]. As in classical physics, ideal measurement is assumed *not* to change the state of the system - the hidden variables would only change in accord with the (as yet unknown) dynamical equations that govern their evolution. The hidden variables are regarded as the *elements of reality* that constitute the fundamental way of describing the system [1]. There may be just a single hidden variable or a set, and the hidden variables could be discrete or continuous - these details do not matter in a general HVT. In the original treatment of Bell [3] the hidden variables uniquely determine the *actual* values that physical quantities would have when measured. However, in a so-called "*fuzzy*" hidden variable theory [82], [83], [84], [85], [15], [6] (see also Section 7.1 of [13]) the values for Ω_A, Ω_B etc are determined *probabilistically* from the hidden variables, the probability functions being *classical* and allow for the hidden variables not being known - just as in classical statistical mechanics, where the unknown (but real) positions and momenta of the classical particles are described via probabilities.

The probabilistic treatment of the hidden variables attempts to replicate the probabilistic nature of quantum theory. For our purposes we will consider only local hidden variable theories (LHV) - this is sufficient to demonstrate key results such as the Bell inequalities. For local hidden variable theories although the hidden variables ξ are *global*, they act locally even for spatially separable subsystems. For particular hidden variables ξ the probability that Ω_A has value λ_i^A will be given by $P_A(i, \xi)$ and the probability that Ω_B has value λ_j^B will be given by $P_B(j, \xi)$, etc. The LHVT joint probability for measurement outcome for Ω_A, Ω_B, etc will be given by

$$P_{AB..}(i, j, ..) = \int d\xi \, P(\xi) \, P_A(i, \xi) P_B(j, \xi)... \tag{34}$$

Here $P(\xi)d\xi$ is the probability that the hidden variables are in the range $d\xi$ around ξ, the HV being assumed continuous - which is not a requirement [3]. The probabilities satisfy the usual sum rules for all outcomes giving unity, thus $\sum_i P_A(i, \xi) = 1$, etc., $\int d\xi \, P(\xi) = 1$. The sub-system probabilities $P_A(i, \xi)$, $P_B(j, \xi)$ etc only depend on the hidden variables ξ. Bell inequalities are constraints derived on the basis of the assumption (34), and if violated therefore falsify all LHV theories.

The *formal similarity* between the hidden variable theory expression for the joint probability (34) and the quantum expression (23) for a separable state is noticable. We could map $\xi \to R$, $P(\xi) \to P_R$, $\int d\xi \Rightarrow \sum_R$, $P_A(i, \xi) \to P_A^R(i)$ and $P_B(j, \xi) \to P_B^R(j)$. The Werner preparation process [5] would then determine the setting for the hidden variables ξ. If a hidden variable theory *underpinned* quantum theory, it follows that the quantum probabilities $P_A^R(i)$ and $P_B^R(j)$ would always be equivalent to hidden variable probabilites $P_A(i, \xi)$ or $P_B(j, \xi)$ for each of the sub-systems (it would not be consistent to only have this apply to one of the sub-systems and not the other). From the expression (34) for the joint probability general hidden variable theory expressions for the mean value $\langle \Omega_A \times \Omega_B \rangle_{HVT}$ for the product of the measurement results for observables Ω_A and Ω_B for subsystems A, B respectively (see (36) below) can be obtained that are analogous to the quantum expression (32) for a separable state. There is of course no independent fully developed classical HVT that can actually predict the $P_A(i, \xi)$, $P_B(j, \xi)$.etc.

However, as we will see both the HVT (see [14] for a proof) and the quantum separable state predictions are consistent with Bell Inequalities, and it therefore requires a quantum entangled state to violate Bell inequalities and to demonstrate failure of the LHV theory model (34). Naturally it follows that such quantum entangled states cannot be described via a LHV theory. Hence the *experimental violation* of Bell inequalities would *also* show that the particular quantum state must be *entangled*.

2.5.1 LHV - Mean Values and Correlation

The actual values that would be assigned to the physical quantities Ω_A, Ω_B etc will depend on the hidden variables but can be taken as the *mean values* of the

possible values λ_i^A, λ_i^A etc. We denote these mean values as $\langle \Omega_A(\xi) \rangle$, $\langle \Omega_B(\xi) \rangle$ etc where

$$\langle \Omega_K(\xi) \rangle = \sum_{\lambda_k^K} \lambda_k^K P_K(k, \xi) \qquad (K = A, B, ...) \tag{35}$$

These expressions my be compared to Eq.(33) for the mean values of physical quantities $\widehat{\Omega}_A$, $\widehat{\Omega}_B$ etc in quantum separable states.

We can then obtain an expression for the mean value in HVT of the physical quantity $\Omega_A \times \Omega_B \times \Omega_C \times ...$, where Ω_A, Ω_B, etc. are physical quantities for the separate sub-systems. This is obtained from Eqs.(34) and (35) and is given by

$$\langle \Omega_A \times \Omega_B \times \Omega_C \times . \rangle_{LHV} = \int d\xi\, P(\xi)\, \langle \Omega_A(\xi) \rangle\, \langle \Omega_B(\xi) \rangle\, \langle \Omega_C(\xi) \rangle .. \tag{36}$$

This may be compared to Eq.(32) for the mean value of the physical quantity $\widehat{\Omega}_A \otimes \widehat{\Omega}_B \otimes \widehat{\Omega}_C \otimes ..$ in quantum separable states.

2.5.2 LHV - GHZ State

The GHZ state [86], [87] is an entangled state of three sub-systems A, B and C, each of which is associated with two quantum states $|+1\rangle$ and $|-1\rangle$. Each sub-system has three physical quantities, which correspond to Pauli spin operators $\widehat{\sigma}_x$, $\widehat{\sigma}_y$ and $\widehat{\sigma}_z$. The quantum states $|+1\rangle$ and $|-1\rangle$ are eigenstates of $\widehat{\sigma}_z$ with eigenvalues $+1$ and -1 respectively. Note that the eigenvalues of the other two Pauli spin operators are also $+1$ and -1. The GHZ state is defined by

$$|\Psi\rangle_{GHZ} = (|+1\rangle_A |+1\rangle_B |+1\rangle_C + |-1\rangle_A |-1\rangle_B |-1\rangle_C)/\sqrt{2} \tag{37}$$

The GHZ state provides a clear example of an entangled quantum state which cannot be described via local hidden variable theory [87], [88]. In a *non-fuzzy* version of the LHV model each of the nine physical quantities σ_x^A, σ_y^A, σ_z^A, σ_x^B, σ_y^B, σ_z^B, σ_x^C, σ_y^C, σ_z^C will be associated with hidden variables that *directly* specify the values $+1$ and -1 that each one of these physical quantities may have. We denote these hidden variables as M_α^K, where $K = A, B, C$ and $\alpha = x, y, z$ and we have $M_\alpha^K = +1$ or -1. With this direct specification of the physical values Eq.(35) just becomes $\langle \sigma_\alpha^K(M^K) \rangle = M_\alpha^K$ and Eq.(36) becomes $\left\langle \sigma_\alpha^A \times \sigma_\beta^B \times \sigma_\gamma^C . \right\rangle_{LHV} = M_\alpha^A M_\beta^B M_\gamma^C$ .We can then derive a contradiction with quantum theory regarding the LHV description of the GHZ state.

Firstly, using the Pauli spin matrices for the $|+1\rangle$ and $|-1\rangle$ basis states

$$[\widehat{\sigma}_x] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad [\widehat{\sigma}_y] = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad [\widehat{\sigma}_x] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{38}$$

it is straightforward to show that the GHZ state satisfies three eigenvalue equations

$$\begin{aligned}
\widehat{\sigma}_x^A \widehat{\sigma}_y^B \widehat{\sigma}_y^C \, |\Psi\rangle_{GHZ} &= (-1) |\Psi\rangle_{GHZ} \\
\widehat{\sigma}_y^A \widehat{\sigma}_x^B \widehat{\sigma}_y^C \, |\Psi\rangle_{GHZ} &= (-1) |\Psi\rangle_{GHZ} \\
\widehat{\sigma}_y^A \widehat{\sigma}_y^B \widehat{\sigma}_x^C \, |\Psi\rangle_{GHZ} &= (-1) |\Psi\rangle_{GHZ}
\end{aligned} \tag{39}$$

Hence in LHV the three quantities $\sigma_x^A \sigma_y^B \sigma_y^C$, $\sigma_y^A \sigma_x^B \sigma_y^C$ and $\sigma_y^A \sigma_y^B \sigma_x^C$ must all have value -1 in the GHZ state, so that as the values for these quantities are just the products of the values for each of the factors we get three equations

$$M_x^A \, M_y^B M_y^C = -1 \quad M_y^A \, M_x^B M_y^C = -1 \quad M_y^A \, M_y^B M_x^C = -1 \qquad (40)$$

Secondly, if we apply all three operators $\widehat{\sigma}_x^A \widehat{\sigma}_x^B \widehat{\sigma}_x^C$ to the GHZ state we find another eigenvalue equation

$$\widehat{\sigma}_x^A \widehat{\sigma}_x^B \widehat{\sigma}_x^C \, |\Psi\rangle_{GHZ} = (+1) \, |\Psi\rangle_{GHZ} \qquad (41)$$

which leads to

$$M_x^A \, M_x^B M_x^C = +1 \qquad (42)$$

However, if we multiply the three equations in Eq.(40) together and use $(M_y^K)^2 = +1$ we find that $M_x^A \, M_x^B M_x^C = -1$, in direct contradiction to the last equation. Thus the assignment of hidden variables for all the physical quantities σ_α^K fails to describe the GHZ state. As we will see in the next SubSection, there are tests involving the violation of Bell Inequalities that are satisfied by some entangled states which allow a demonstration of the failure of more general local LHV theories, even allowing for correlations that are less than ideal.

The assumption of non-fuzzy LHV theories is not essential for the GHZ arguement in the case of the ideal GHZ state (37). This is because one may use the correlations of (39) to establish a *precise* prediction of one of the spins at A, by measuring the spins at the other two locations. The assumption of local realism (on which the LHV theory is based) then establishes a precise value for the hidden variable [86], [87], [88]. In a more realistic scenario where the GHZ correlations are not perfect, the "elements of reality" established this way becme fuzzy, and in that case Mermin's Bell inequality [88] can be used to establish a contradiction with LHV models.

2.6   Paradoxes

The EPR and Schrodinger Cat paradoxes figured prominently in early discussions about entanglement. Both paradoxes involve *composite systems* and the consideration of quantum states which are entangled Both these paradoxes reflect the conflict between *quantum theory*, in which the values for physical quantities only take on definite values when measurement occurs and *classical theory*, in which the values for physical quantities always exist even when measurement is not involved. The latter viewpoint is referred to as *realism*. Quantum theory is also probabalistic, so although the possible outcomes for measuring a physical quantity can be determined prior to measurement, the actual outcome in a given quantum state for any measurement is only known in terms of a *probability*. However, from the realist viewpoint, quantum theory is *incomplete* and a *future theory* based around *hidden variables* would determine the actual values of the physical quantities, as well as the quantum probabilities that particular values will be found via measurement.

Whilst the EPR and Schrodinger Cat paradoxes are of historical interest and have provoked much debate, it was the formulation of the *Bell inequalities* (which are described in the next SubSection 2.7) and the conditions under which they could be violated that provided the first clear case of where the predictions of quantum theory could differ from those of hidden variable theories. It then became possible to carry out actual experiments to distinguish these two fundamentally different theories. The actual experimental evidence is consistent with quantum theory and (apart from a small number of remaining loopholes) rules out local hidden variable theories.

2.6.1 EPR Paradox

In the original version of the EPR paradox, Einstein et al [1] considered a two-particle system A, B in which the particles were associated with *positions* $\widehat{x}_A$, $\widehat{x}_B$ and *momenta* $\widehat{p}_A$, $\widehat{p}_B$. They envisaged a quantum state in which the pairs of physical quantities $\widehat{x}_A$, $\widehat{x}_B$ or $\widehat{p}_A$, $\widehat{p}_B$ had highly *correlated* values - measured or otherwise. To be specific, one may consider a simultaneous eigenstate $|\Phi\rangle$ of the two commuting operators $\widehat{x}_A - \widehat{x}_B$ and $\widehat{p}_A + \widehat{p}_B$, where $(\widehat{x}_A - \widehat{x}_B)|\Phi\rangle = 2x|\Phi\rangle$ and $(\widehat{p}_A + \widehat{p}_B)|\Phi\rangle = 0|\Phi\rangle$. This state is an example of an entangled state, as may be seen if it is expanded in terms of position eigenstates $|x_A x_B\rangle$. If the system is in state $|\Phi\rangle$ then from standard quantum theory if A had a mean momentum p then B would have a mean momentum $-p$. Alternatively, if A had a mean position x then B would have a mean position $-x$. Then if the eigenvalue $2x$ is very large so that the two particles will be well-separated (in quantum theory their spatial wave functions would be localised in separate spatial regions) it follows that if the position of B was measured then the position of A would be immediately known, even if the particles were light years apart. On the other hand, if the momentum of B was measured instead, then the momentum of A would immediately be known. From the realist point of view both A and B always have definite positions and momenta, even if these are not known, so all these measurements do is reveal these (hidden) values. It would seem then that measurements of position and momentum on particle B *could* lead to a knowledge of the position and momentum at a far distant particle A, perhaps with an accuracy that would violate the Heisenberg Uncertainty Principle (HUP). As we will see, this is not the case when quantum theory is applied correctly. However, what Einstein et al pointed out as being particularly strange was that the choice of whether the momentum or position of B was measured (and found to have a definite value) would instantly determine which of the position or momentum of A would then have a definite value - even if A and B were separated by such a large distance that no signal could have been passed from B to A regarding which quantity was measured. Einstein referred to this as "spooky action at a distance" to highlight the strangeness of what came to be referred to as entangled states. Thus a somewhat paradoxical situation would seem to arise. Einstein stated that this did not demonstrate that quantum theory was wrong, only that it was incomplete.

The EPR argument assumes *local realism*, to justify that the posibility of

an exact *prediction* of the postion of the far-away particle A (based on the measurement of the position for the particle B) implies the realist viewpoint that the position of particle A was predetermined. The same argument applies to the momentum of particle A, and hence EPR conclude that *both* the position and momentum of particle A are precisely predetermined - in conflict with the Heisenberg Uncertainty Principle derived from quantum mechanics. Since the argument is based on the assumption of local realism, the modern interpretation of the EPR analysis is that it reveals (for the appropriate entangled state) the inconsistency of local realism with the completeness of quantum mechanics.

Discussions of the *EPR paradox* [1] in terms of hidden variable theories has been given by numerous authors (see [85], [13], [14], [15] for example). The papers and reviews by Reid et al [76], [85], [15], give a full account taking into consideration the "fuzzy" version of *local* HVT (LHV) and determining the predictions for the conditional variances for x_A and p_A based both on separable quantum states and states described via local HVT. This treatment successfully quantifies the somewhat qualitative considerations described in the previous paragraph. If the position for particle B is measured and the result is x, then the original density operator $\widehat{\rho}$ for the two particle system is changed into the conditional density operator $\widehat{\rho}_{cond}(\widehat{x}_B, x) = \widehat{\Pi}_x^B \widehat{\rho} \widehat{\Pi}_x^B / Tr(\widehat{\Pi}_x^B \widehat{\rho})$, where $\widehat{\Pi}_x^B = (|x\rangle \langle x|)_B$ is the projector onto the eigenvector $|x\rangle_B$ (the eigenvalues x are assumed for simplicity to form a quasi-continuum). Similarly, if the momentum for particle B is measured and the result is p, then the original density operator $\widehat{\rho}$ for the two particle system is changed into the conditional density operator $\widehat{\rho}_{cond}(\widehat{p}_B, p) = \widehat{\Pi}_p^B \widehat{\rho} \widehat{\Pi}_p^B / Tr(\widehat{\Pi}_p^B \widehat{\rho})$, where $\widehat{\Pi}_p^B = (|p\rangle \langle p|)_B$ is the projector onto the eigenvector $|p\rangle_B$ (the eigenvalues p are assumed for simplicity to form a quasi-continuum). Here we outline the discussion based on quantum separable states. Conditional variances for position and momentum for sub-system A are considered based on measurements for sub-system B of position. It is shown that for these conditional variances the Heisenberg uncertainty principle still applies. The same conclusion is obtained if the measurements on sub-system B had been the momentum. As the experimenter on sub-system A could not know whether the measurement on sub-system B was on position or momentum, the action at a distance feature of quantum entanglement is confirmed.

The question is whether the conditional variances $\left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B}$ for measuring $\widehat{x}_A$ for sub-system A having measured $\widehat{x}_B$ for sub-system B, and $\left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B}$ for measuring $\widehat{p}_A$ for sub-system A having measured $\widehat{p}_B$ for sub-system B violate the Heisenberg Uncertainty Principle [76]

$$\left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B} \left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B} < \frac{1}{4}\hbar^2 \tag{43}$$

where the measurements on sub-system B are left unrecorded. If this inequality holds we have an *EPR violation*.

For *separable states* the *conditional probability* that measurement of $\widehat{x}_A$ on sub-system A leads to eigenvalue x_A given that measurement of $\widehat{x}_B$ on sub-

system B leads to eigenvalue x_B is obtained from Eq.(30) as

$$P(\widehat{x}_A, x_A|\widehat{x}_B, x_B) = \sum_R P_R \, P_A^R(\widehat{x}_A, x_A) P_B^R(\widehat{x}_B, x_B) / \sum_R P_R \, P_B^R(\widehat{x}_B, x_B) \quad (44)$$

where

$$P_A^R(\widehat{x}_A, x_A) = Tr_A(\widehat{\Pi}_{x_A}^A \widehat{\rho}_R^A) \qquad P_B^R(\widehat{x}_B, x_B) = Tr_B(\widehat{\Pi}_{x_B}^B \widehat{\rho}_R^B) \qquad (45)$$

are the probabilities for position measurements in the separate sub-systems. The probability that measurement of $\widehat{x}_B$ on sub-system B leads to eigenvalue x_B is

$$P(\widehat{x}_B, x_B) = \sum_R P_R \, P_B^R(\widehat{x}_B, x_B) \qquad (46)$$

The *mean* result for measurement of $\widehat{x}_A$ for this *conditional* measurement is from Eq.(19)

$$\begin{aligned}
\langle \widehat{x}_A \rangle_{\widehat{x}_B, x_B} &= \sum_{x_A} x_A \, P(\widehat{x}_A, x_A|\widehat{x}_B, x_B) \\
&= \sum_R P_R \, \langle \widehat{x}_A \rangle_R \, P_B^R(\widehat{x}_B, x_B) / P(\widehat{x}_B, x_B) \qquad (47)
\end{aligned}$$

where

$$\langle \widehat{x}_A \rangle_R = \sum_{x_A} x_A P_A^R(\widehat{x}_A, x_A) \qquad (48)$$

is the *mean* result for measurement of $\widehat{x}_A$ when the sub-system is in state $\widehat{\rho}_R^A$.

The *conditional variance* for measurement of $\widehat{x}_A$ for the conditional measurement of $\widehat{x}_B$ on sub-system B which led to eigenvalue x_B is from Eq.(20)

$$\begin{aligned}
\langle \Delta \widehat{x}_A^2 \rangle_{\widehat{x}_B, x_B} &= \sum_{x_A} (x_A - \langle \widehat{x}_A \rangle_{\widehat{x}_B, x_B})^2 \, P(\widehat{x}_A, x_A|\widehat{x}_B, x_B) \\
&= \sum_R P_R \, \langle \Delta \widehat{x}_A^2 \rangle_{\widehat{x}_B, x_B}^R \, P_B^R(\widehat{x}_B, x_B) / P(\widehat{x}_B, x_B) \qquad (49)
\end{aligned}$$

where

$$\langle \Delta \widehat{x}_A^2 \rangle_{\widehat{x}_B, x_B}^R = \sum_{x_A} (x_A - \langle \widehat{x}_A \rangle_{\widehat{x}_B, x_B})^2 \, P_A^R(\widehat{x}_A, x_A)$$

is a variance for measurement of $\widehat{x}_A$ for when the sub-system is in state $\widehat{\rho}_R^A$ but now with the fluctuation about the mean $\langle \widehat{x}_A \rangle_{\widehat{x}_B, x_B}$ for measurements conditional on measuring $\widehat{x}_B$.

However, for each sub-system state R the quantity $\langle \Delta \widehat{x}_A^2 \rangle_{\widehat{x}_B, x_B}^R$ is *minimised* if $\langle \widehat{x}_A \rangle_{\widehat{x}_B, x_B}$ is replaced by the unconditioned mean $\langle \widehat{x}_A \rangle_R$ just determined from $\widehat{\rho}_R^A$. Thus we have an inequality

$$\langle \Delta \widehat{x}_A^2 \rangle_{\widehat{x}_B, x_B}^R \geq \langle \Delta \widehat{x}_A^2 \rangle^R \qquad (50)$$

where

$$\left\langle \Delta \widehat{x}_A^2 \right\rangle^R = \sum_{x_A} (x_A - \langle \widehat{x}_A \rangle)^2 \, P_A^R(\widehat{x}_A, x_A) \tag{51}$$

is the *normal variance* for measurement of $\widehat{x}_A$ for when the sub-system is in state $\widehat{\rho}_R^A$.

Now if the measurements of $\widehat{x}_B$ are *unrecorded* - as would be the case from the point of view of the experimenter on spatially well-separated sub-system A when measurements on this sub-system take place at the same time - then the *conditioned variance* is

$$\begin{aligned}
\left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B} &= \sum_{x_B} \left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B, x_B} P(\widehat{x}_B, x_B) \\
&= \sum_{x_B} \sum_R P_R \left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B, x_B}^R P_B^R(\widehat{x}_B, x_B)
\end{aligned} \tag{52}$$

which in view of inequality (50) satisfies

$$\begin{aligned}
\left\langle \Delta \widehat{x}_A^2 \right\rangle_{\widehat{x}_B} &\geq \sum_{x_B} \sum_R P_R \left\langle \Delta \widehat{x}_A^2 \right\rangle^R P_B^R(\widehat{x}_B, x_B) \\
&= \sum_R P_R \left\langle \Delta \widehat{x}_A^2 \right\rangle^R
\end{aligned} \tag{53}$$

using $\sum_{x_B} P_B^R(\widehat{x}_B, x_B) = 1$. Thus the variance for measurement of position $\widehat{x}_A$ conditioned on unrecorded measurements for position $\widehat{x}_B$ satisfies an inequality that only depends on the variances for measurements of $\widehat{x}_A$ in the possible sub-system A states $\widehat{\rho}_R^A$.

Now exactly the same treatment can be carried out for the variance of *momentum* $\widehat{p}_A$ also conditioned on unrecorded measurements of measurements for momentum $\widehat{x}_B$. Details are given in Appendix 6. We have with

$$\begin{aligned}
\left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B} &= \sum_{p_B} \left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B, p_B} P(\widehat{p}_B, p_B) \\
\left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B, p_B} &= \sum_{p_A} (p_A - \langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B})^2 \, P(\widehat{p}_A, p_A | \widehat{p}_B, p_B) \\
\langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B} &= \sum_{p_A} p_A \, P(\widehat{p}_A, p_A | \widehat{p}_B, p_B)
\end{aligned}$$

the inequality

$$\left\langle \Delta \widehat{p}_A^2 \right\rangle_{\widehat{p}_B} \geq \sum_R P_R \left\langle \Delta \widehat{p}_A^2 \right\rangle^R \tag{54}$$

with

$$\left\langle \Delta \widehat{p}_A^2 \right\rangle^R = \sum_{p_A} (p_A - \langle \widehat{p}_A \rangle)^2 \, P_A^R(\widehat{p}_A, p_A) \tag{55}$$

is the normal variance for measurement of $\widehat{p}_A$ for when the sub-system is in state $\widehat{\rho}_R^A$.

We now multiply the two conditional variances, which it is important to note were associated with two *different conditioned states* based on two *different* measurements - position and momentum - carried out on sub-system B.

$$\left\langle \Delta\widehat{x}_A^2\right\rangle_{\widehat{x}_B} \left\langle \Delta\widehat{p}_A^2\right\rangle_{\widehat{p}_B} \geq \sum_R P_R \left\langle \Delta\widehat{x}_A^2\right\rangle^R \sum_S P_S \left\langle \Delta\widehat{p}_A^2\right\rangle^S \qquad (56)$$

However, from the general inequality in Eq.(177)

$$\sum_R P_R\, C_R \sum_R P_R\, D_R \geq \left(\sum_R P_R\, \sqrt{C_R D_R}\right)^2 \qquad (57)$$

we then have

$$\begin{aligned}
\left\langle \Delta\widehat{x}_A^2\right\rangle_{\widehat{x}_B} \left\langle \Delta\widehat{p}_A^2\right\rangle_{\widehat{p}_B} &\geq& \left(\sum_R P_R \sqrt{\left\langle \Delta\widehat{x}_A^2\right\rangle^R \left\langle \Delta\widehat{p}_A^2\right\rangle^R}\right)^2 \\
&=& \left(\sum_R P_R \sqrt{\left\langle \Delta\widehat{x}_A^2\right\rangle^R} \times \sqrt{\left\langle \Delta\widehat{p}_A^2\right\rangle^R}\right)^2 \qquad (58)
\end{aligned}$$

But we know from the HUP that for *any* given state $\widehat{\rho}_R^A$ that $\left\langle \Delta\widehat{x}_A^2\right\rangle^R \left\langle \Delta\widehat{p}_A^2\right\rangle^R \geq \frac{1}{4}\hbar^2$, so for the conditioned variances associated with a *separable* state

$$\left\langle \Delta\widehat{x}_A^2\right\rangle_{\widehat{x}_B} \left\langle \Delta\widehat{p}_A^2\right\rangle_{\widehat{p}_B} \geq \frac{1}{4}\hbar^2 \qquad (59)$$

showing that for a separable state the conditioned variances $\left\langle \Delta\widehat{x}_A^2\right\rangle_{\widehat{x}_B}$ and $\left\langle \Delta\widehat{p}_A^2\right\rangle_{\widehat{p}_B}$ still satisfy the HUP. It is important to note that these variances were associated with two different conditioned states based on two different measurements - position and momentum - carried out on sub-system B, the results of which the observer for sub-system A would be unaware of. Thus if the EPR violations as defined in Eq.(43) are to occur then the state must be entangled. Progress towards experimental confirmation of EPR violations is reviewed in Refs. [15], [6].

In [85] an analogous treatment based on *local hidden variable theory* (LHV) also shows that the HUP is satisfied for the conditioned variances. The details of this treatment will not be given here, but the formal similarity of expressions for conditional probabilities in LHV theories and for separable states indicates the steps involved.

The EPR paradox is not confined to position and momentum measurements on two sub-systems. A related paradox [89] occurs in the case of measurements on *spin components* $\widehat{S}_{\alpha 1}$ and $\widehat{S}_{\alpha 2}$ - with $\alpha = x, y, z$ - associated with two sub-systems 1 and 2. The spin operators also satisfy non-zero commutation rules (see paper II for details)

$$[\widehat{S}_{\alpha 1}, \widehat{S}_{\beta 1}] = i\widehat{S}_{\gamma 1} \qquad [\widehat{S}_{\alpha 2}, \widehat{S}_{\beta 2}] = i\widehat{S}_{\gamma 2} \qquad (60)$$

where α, β, γ are x, y, z in cyclic order. Different spin components for each sub-system do not have simultaneous precise measurements leading to Heisenberg Uncertainty principle relations involving the *variances* and *mean* values

$$\left\langle \Delta \widehat{S}_{\alpha1}^2 \right\rangle \left\langle \Delta \widehat{S}_{\beta1}^2 \right\rangle \geq \frac{1}{4} |\left\langle \widehat{S}_{\gamma1} \right\rangle|^2 \qquad \left\langle \Delta \widehat{S}_{\alpha2}^2 \right\rangle \left\langle \Delta \widehat{S}_{\beta2}^2 \right\rangle \geq \frac{1}{4} |\left\langle \widehat{S}_{\gamma2} \right\rangle|^2 \qquad (61)$$

As in the case of position and momentum a special state of the combined system has interesting features. For the case where the spin quantum number of each sub-system is $1/2$ the measured values for *any* spin component of either system is either $+1/2$ or $-1/2$. In terms of eigenstates for $\widehat{S}_{x1}$ and $\widehat{S}_{x2}$ we consider the state

$$\left| \Psi^- \right\rangle = \frac{1}{\sqrt{2}} \left(|x, +\rangle_1 \otimes |x, -\rangle_2 - |x, -\rangle_1 \otimes |x, +\rangle_2 \right) \qquad (62)$$

This is actually one of the Bell states. In this form it shows that measurements of the x components of the spins are perfectly correlated, so that for example if the measurement of $\widehat{S}_{x2}$ for sub-system 2 results in the value $-1/2$, then a subsequent measurement of $\widehat{S}_{x1}$ for sub-system 1 must result in the value $+1/2$. However, the same state can be expressed in terms of eigenstates for $\widehat{S}_{y1}$ and $\widehat{S}_{y2}$ as

$$\left| \Psi^- \right\rangle = \frac{1}{\sqrt{2}} \left(|y, +\rangle_1 \otimes |y, -\rangle_2 - |y, -\rangle_1 \otimes |y, +\rangle_2 \right) \qquad (63)$$

and analogous statements regarding measurement correlations apply if the measurements were for $\widehat{S}_{y2}$ on sub-system 2 with a subsequent measurement of $\widehat{S}_{y1}$ on sub-system 1. If the two sub-systems were well-separated it might be expected that first measuring $\widehat{S}_{x2}$ for sub-system 2 would determine the result of measuring $\widehat{S}_{x1}$ for sub-system 1, and then measuring $\widehat{S}_{y2}$ for sub-system 2 would determine the result of measuring $\widehat{S}_{y1}$ for sub-system 1 - and as the second ($\widehat{S}_{y2}$) measurement on far distant sub-system 2 should not affect the former measurement on sub-system 1 this would appear to result in precise measured values for $\widehat{S}_{x1}$ and $\widehat{S}_{y1}$ on sub-system 1, which conflicts with the Heisenberg Uncertainty principle requirement that $\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle \geq \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle|^2$.

However we can consider the variances for $\widehat{S}_{x1}$ and $\widehat{S}_{y1}$ which are conditional on measurements for $\widehat{S}_{x2}$ and $\widehat{S}_{y2}$ for sub-system 2 and show that for *separable* states of the two sub-systems we have

$$\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} \geq \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle|^2 \qquad (64)$$

Thus if we find that

$$\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} < \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle|^2 \qquad (65)$$

then we have an example of a *spin EPR violation.* Such a violation requires that the quantum state is *entangled.* The derivation of the result (64) for separable states is set out in Appendix 8.

An effect related to the EPR paradox is *EPR Steering*. As we have seen, the measurement of the position for particle B changes the density operator and consequently the probability distributions for measurements on particle A will now be determined from the conditional probabilities, such as $P_{AB}(\widehat{x}_A, x_A | \widehat{x}_B, x_B)$ or $P_{AB}(\widehat{p}_A, p_A | \widehat{x}_B, x_B)$.Thus measurements on B are said to steer the results for measurements on A. Steering will of course only apply if the measurement results for $\widehat{x}_B$ are *recorded*, and not discarded. A discussion of EPR Steering (see [15], [26]) is beyond the scope of this article.

2.6.2   Schrodinger Cat Paradox

The Schrodinger Cat Paradox [2], [90] relates to composite systems where one sub-system (the cat) is macroscopic and the other sub-system is microscopic (the radioactive atom). The paradox is a clear consequence of quantum theory allowing the existence of entangled states. Schrodinger envisaged a state in which an alive cat and an undecayed atom existed at an initial time, and because the decayed atom would be associated with a dead cat, the system after a time corresponding to the half-life for radioactive decay would be described in quantum theory via the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|e\rangle_{Atom} |Alive\rangle_{Cat} + |g\rangle_{Atom} |Dead\rangle_{Cat}) \tag{66}$$

in an obvious notation. The quantum state defined by (66) represents the knowledge that an observer outside the box would have about the combined atom-cat system one hour after the live cat was placed in the box along with an undecayed atom. The combined system is in an enclosed box, and opening the box and observing what is inside constitutes a measurement on the system. According to quantum theory if the box was opened at this time there would be a probability of $1/2$ of finding the atom undecayed and the cat alive, with the same probability for finding a decayed atom and a dead cat. From the realist viewpoint the cat should be *either* dead *or* it should be alive *irrespective* of whether the box is opened or not, and it was regarded as a paradox that in the quantum theory description of the state prior to measurement the cat is in some sense *both* dead *and* alive. This paradox is made worse because the cat is a macroscopic system - how could a cat be either dead or alive at the same time, it must be one or the other? From the quantum point of view in which the actual values of physical quantities *only* appear when measurement occurs, the Schrodinger cat presents no paradox. The two possible values signifying the health of the cat are "alive" and "dead", and these values are found with a probability of $1/2$ when measurement takes place on opening the box, and this would entirely explain the results if such an experiment were to be performed. There is of course no paradox if the quantum state is only considered to represent the observer's information about what is inside the box. If the box is closed then at one half life after the cat was put into the box, the state vector (66) enables the outside observer to correctly assess the probability that the cat will be alive is $1/2$. If the box is then opened and the cat is found to be dead, then

the observer's information changes and the state vector for the cat-atom system is now

$$\left|\Psi^{'}\right\rangle = |g\rangle_{Atom}\,|Dead\rangle_{Cat} \tag{67}$$

In this interpretation of quantum states, the notion of there being some sort of underlying reality that exists *prior* to measurement is rejected. It is only this notion that such a reality must exist - perhaps described via hidden variables - that leads to the paradox. EPR paradoxes can also be constructed from the entangled state (66), as outlined in Refs. [91], [92].

In recent times, experiments based on a *Rydberg atom* in a *microwave cavity* [93] involving states such as (66) have been performed showing that entanglement can occur between macroscopic and microscopic systems, and it is even possible to prepare states analogous to $\frac{1}{\sqrt{2}}(|Alive\rangle_{Cat}+|Dead\rangle_{Cat})$ in the macroscopic system itself. In such experiments the different macroscopic states are large amplitude coherent states of the cavity mode. Coherent states are possible for *microwave photons* as they are created from *classical currents* wiith *well-defined phases.* A coherent superposition of an alive and dead cat within the cat sub-system itself can be *created* by measurement. The entangled state in (66) can also be written as

$$
\begin{aligned}
|\Psi\rangle &= \frac{1}{\sqrt{2}}\{\frac{1}{\sqrt{2}}\left(|e\rangle_{Atom}+|g\rangle_{Atom}\right)\frac{1}{\sqrt{2}}(|Alive\rangle_{Cat}+|Dead\rangle_{Cat}) \\
&\quad +\frac{1}{\sqrt{2}}\left(|e\rangle_{Atom}-|g\rangle_{Atom}\right)\frac{1}{\sqrt{2}}(|Alive\rangle_{Cat}-|Dead\rangle_{Cat})\}
\end{aligned} \tag{68}
$$

so that measurement on the atom for an observable in which the superposition states $\frac{1}{\sqrt{2}}\left(|e\rangle_{Atom}\pm|g\rangle_{Atom}\right)$ are the eigenstates for this observable would result in the cat *then* being in the corresponding.macroscopic superposition states $\frac{1}{\sqrt{2}}(|Alive\rangle_{Cat}\pm|Dead\rangle_{Cat})$ of an alive and dead cat.

2.7   Bell Inequalities

Violations of Bell's Inequalities represent situations where neither hidden variable theory nor quantum theory based on separable states can account for the result, and therefore provide a clear case where an entangled quantum state is involved.

2.7.1   Local Hidden Variable Theory Result

A key feature of entangled states is that they are associated with *violations of Bell inequalities* [3] and hence can exhibit this particular *non-classical* feature. The Bell inequalities arise in attempts to restore a *classical* interpretation of quantum thory via hidden variable treatments, where actual values are assigned to all measureable quantities - including those which in quantum theory are associated with non-commuting Hermitian operators. In this case we consider two different physical quantities Ω_A for sub-system A, which are listed A_1, A_2, etc , and two Ω_B for sub-system B, which are listed B_1, B_2, etc. The

corresponding quantum Hermitian operators $\widehat{\Omega}_A$, $\widehat{\Omega}_B$, etc are $\widehat{A}_1$, $\widehat{A}_2$ and, $\widehat{B}_1$, $\widehat{B}_2$. The Bell inequalities involve the *mean value* $\langle A_i \times B_j \rangle_{HVT}$ of the product of observables A_i and B_j for subsystems A, B respectively, for which there are two possible measured values, $+1$ and -1. For simplicity we consider a local HVT. In a local hidden variable theory.(LHV) we see using (34) that the mean values $\langle A_i \times B_j \rangle_{LHV}$ are given by

$$\langle A_i \times B_j \rangle_{LHV} = \int d\xi \, P(\xi) \, \langle A_i(\xi) \rangle \, \langle B_j(\xi) \rangle \tag{69}$$

where $\langle A_i(\xi) \rangle$ and $\langle B_j(\xi) \rangle$ (as in Eq.(35)) are the values are assigned to A_i and B_j when the hidden variables are ξ, and $P(\xi)$ is the hidden variable probability distribution function. If the corresponding quantum Hermitian operators are such that their eigenvalues are $+1$ and -1 - as in the case of Pauli spin operators - then the only possible values for $\langle A_i(\xi) \rangle$ and $\langle B_j(\xi) \rangle$ are betweem $+1$ and -1, since HVT does not conflict with quantum theory regarding allowed values for physical quantities. However, local hidden variable theory predicts certain inequalities for the mean values of products of physical quantities for the two sub-systems.

The form given by Clauser et al [82] for *Bell's inequality* is

$$|S| \leq 2 \tag{70}$$

where

$$S = \langle A_1 \times B_1 \rangle_{LHV} + \langle A_1 \times B_2 \rangle_{LHV} + \langle A_2 \times B_1 \rangle_{LHV} - \langle A_2 \times B_2 \rangle_{LHV} \tag{71}$$

The minus sign can actually be attached to any one of the four terms.

Following the proof of the Bell inequalities in [14] we have

$$
\begin{aligned}
\langle A_2 \times B_1 \rangle_{LHV} - \langle A_2 \times B_2 \rangle_{LHV} &= \int d\xi \, P(\xi) \left(\langle A_2(\xi) \rangle \, \langle B_1(\xi) \rangle - \langle A_2(\xi) \rangle \, \langle B_2(\xi) \rangle \right) \\
&= \int d\xi \, P(\xi) \left(\langle A_2(\xi) \rangle \, \langle B_1(\xi) \rangle \, (1 \pm \langle A_1(\xi) \rangle \, \langle B_2(\xi) \rangle) \right) \\
&\quad - \int d\xi \, P(\xi) \left(\langle A_2(\xi) \rangle \, \langle B_2(\xi) \rangle \, (1 \pm \langle A_1(\xi) \rangle \, \langle B_1(\xi) \rangle) \right)
\end{aligned}
\tag{72}
$$

Now all the quantities $\langle A_i(\xi) \rangle$, $\langle B_j(\xi) \rangle$ are bounded by $+1$ or -1, so the expressions $(1 \pm \langle A_1(\xi) \rangle \, \langle B_2(\xi) \rangle)$ and $(1 \pm \langle A_1(\xi) \rangle \, \langle B_1(\xi) \rangle)$ are never negative.

Taking the modulus of the left side leads to an equality

$$\left|\langle A_2 \times B_1\rangle_{LHV} - \langle A_2 \times B_2\rangle_{LHV}\right|$$

$$\leq \int d\xi\, P(\xi)\left(\left|\langle A_2(\xi)\rangle\right|\left|\langle B_1(\xi)\rangle\right|\left(1 \pm \langle A_1(\xi)\rangle\, \langle B_2(\xi)\rangle\right)\right)$$

$$+ \int d\xi\, P(\xi)\left(\left|\langle A_2(\xi)\rangle\right|\left|\langle B_2(\xi)\rangle\right|\left(1 \pm \langle A_1(\xi)\rangle\, \langle B_1(\xi)\rangle\right)\right)$$

$$\leq \int d\xi\, P(\xi)\left(1 \pm \langle A_1(\xi)\rangle\, \langle B_2(\xi)\rangle\right) + \int d\xi\, P(\xi)\left(1 \pm \langle A_1(\xi)\rangle\, \langle B_1(\xi)\rangle\right)$$

$$= 2 \pm \left(\int d\xi\, P(\xi)\, \langle A_1(\xi)\rangle\, \langle B_2(\xi)\rangle + \int d\xi\, P(\xi)\, \langle A_1(\xi)\rangle\, \langle B_1(\xi)\rangle\right)$$

$$= 2 \pm \left(\langle A_1 \times B_2\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV}\right) \tag{73}$$

where we have used the results that $\left|\langle A_2(\xi)\rangle\right|, \left|\langle B_1(\xi)\rangle\right|$ and $\left|\langle B_2(\xi)\rangle\right|$ are all less than unity and that $\int d\xi\, P(\xi) = 1$. Hence since $\left|\langle A_1 \times B_2\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV}\right| = +(\langle A_1 \times B_2\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV})$ or $-(\langle A_1 \times B_2\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV})$ we have

$$\left|\langle A_2 \times B_1\rangle_{LHV} - \langle A_2 \times B_2\rangle_{LHV}\right| \pm \left|\langle A_1 \times B_2\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV}\right| \leq 2 \tag{74}$$

But since $|X - Y| \leq |X| + |Y|$ we see that from the + version of the last inequality that

$$\left|\langle A_2 \times B_1\rangle_{LHV} - \langle A_2 \times B_2\rangle_{LHV} + \langle A_1 \times B_2\rangle_{HVT} + \langle A_1 \times B_1\rangle_{HVT}\right| \leq 2 \tag{75}$$

This is a Bell inequality. Interchanging $A_2 \leftrightarrow A_1$ and repeating the derivation gives $\left|\langle A_1 \times B_1\rangle_{HVT} - \langle A_1 \times B_2\rangle_{LHV} + \langle A_2 \times B_2\rangle_{LHV} + \langle A_2 \times B_1\rangle_{LHV}\right| \leq 2$, which is another Bell inequality. Interchanging $B_1 \leftrightarrow B_2$ and repeating the derivation gives $\left|\langle A_2 \times B_2\rangle_{LHV} - \langle A_2 \times B_1\rangle_{LHV} + \langle A_1 \times B_1\rangle_{LHV} + \langle A_1 \times B_2\rangle_{LHV}\right| \leq 2$, and interchanging $A_2 \leftrightarrow A_1$ and $B_1 \leftrightarrow B_2$ and repeating the derivation gives $\left|\langle A_1 \times B_2\rangle_{LHV} - \langle A_1 \times B_1\rangle_{LHV} + \langle A_2 \times B_1\rangle_{LHV} + \langle A_2 \times B_2\rangle_{LHV}\right| \leq 2$. Thus the minus sign can be attached to any one of the four terms. An example of an entangled state that violates the Bell inequality is given in SubSection 2.7.3.

2.7.2 Non-Entangled State Result

It can be shown that the Bell inequalities *also always* occur for *non-entangled* states (see Section 7.3 of the book by Vedral [13]). For Bell's inequalities we consider Hermitian operators $\widehat{A}_i$ and $\widehat{B}_j$ for subsystems A, B respectively, for which there are two eigenvalues $+1$ and -1, where examples of the operators are given by the components $\widehat{A}_i = a_i \cdot \widehat{\sigma}_A$ and $\widehat{B}_j = b_j \cdot \widehat{\sigma}_B$ of Pauli spin operators $\widehat{\sigma}_A$ and $\widehat{\sigma}_B$ along directions with unit vectors a_i and b_j. The corresponding quantum theory quantity for the Bell inequality is

$$S = E(\widehat{A}_1 \otimes \widehat{B}_1) + E(\widehat{A}_1 \otimes \widehat{B}_2) + E(\widehat{A}_2 \otimes \widehat{B}_1) - E(\widehat{A}_2 \otimes \widehat{B}_2) \tag{76}$$

where in quantum theory the mean value is given by $E(\widehat{A}_i \otimes \widehat{B}_j) = \left\langle \widehat{A}_i \otimes \widehat{B}_j \right\rangle = Tr(\widehat{\rho}\,\widehat{A}_i \otimes \widehat{B}_j)$. For the general bipartite non-entangled state given by 3 it is easy to show that

$$S = \sum_R P_R \left(\left\langle \widehat{A}_1 \right\rangle_R^A \left\langle \widehat{B}_1 + \widehat{B}_2 \right\rangle_R^B + \left\langle \widehat{A}_2 \right\rangle_R^A \left\langle \widehat{B}_1 - \widehat{B}_2 \right\rangle_R^B \right) \qquad (77)$$

where $\left\langle \widehat{A}_i \right\rangle_R^A = Tr(\widehat{A}_i\, \widehat{\rho}_R^A)$ and $\left\langle \widehat{B}_j \right\rangle_R^B = Tr(\widehat{B}_j\, \widehat{\rho}_R^B)$ are the expectation values of $\widehat{A}_i$ and $\widehat{B}_j$ for the sub-systems A, B in states $\widehat{\rho}_R^A$ and $\widehat{\rho}_R^B$ respectively. Now $\left\langle \widehat{A}_i \right\rangle_R^A$ and $\left\langle \widehat{B}_j \right\rangle_R^B$ must lie in the range -1 to $+1$, so that $\left\langle \widehat{B}_1 \pm \widehat{B}_2 \right\rangle_R^B$ must each lie in the range -2 to $+2$. Hence

$$
\begin{aligned}
|S| &\leq \sum_R P_R \left(|\left\langle \widehat{A}_1 \right\rangle_R^A|\,|\left\langle \widehat{B}_1 + \widehat{B}_2 \right\rangle_R^B| + |\left\langle \widehat{A}_2 \right\rangle_R^A|\,|\left\langle \widehat{B}_1 - \widehat{B}_2 \right\rangle_R^B| \right) \\
&\leq \sum_R P_R \left(|\left\langle \widehat{B}_1 + \widehat{B}_2 \right\rangle_R^B| + |\left\langle \widehat{B}_1 - \widehat{B}_2 \right\rangle_R^B| \right) \\
&\leq 2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (78)
\end{aligned}
$$

since to obtain $|\left\langle \widehat{B}_1 + \widehat{B}_2 \right\rangle_R^B| = 2$ requires $\left\langle \widehat{B}_1 \right\rangle_R^B = \left\langle \widehat{B}_2 \right\rangle_R^B = \pm 1$ and then $|\left\langle \widehat{B}_1 - \widehat{B}_2 \right\rangle_R^B| = |\left\langle \widehat{B}_1 \right\rangle_R^B - \left\langle \widehat{B}_2 \right\rangle_R^B| = 0$, or to obtain $|\left\langle \widehat{B}_1 - \widehat{B}_2 \right\rangle_R^B| = 2$ requires $\left\langle \widehat{B}_1 \right\rangle_R^B = -\left\langle \widehat{B}_2 \right\rangle_R^B = \pm 1$ and then $|\left\langle \widehat{B}_1 + \widehat{B}_2 \right\rangle_R^B| = |\left\langle \widehat{B}_1 \right\rangle_R^B + \left\langle \widehat{B}_2 \right\rangle_R^B| = 0$.

2.7.3 Bell Inequality Violation and Entanglement

It follows that for a general two mode non-entangled state $|S|$ cannot violate the Bell inequality upper bound of 2. Thus, the violation of Bell inequalities proves that the quantum state must be entangled for the sub-systems involved, so Bell inequality violations are a test of entanglement. For entangled states such as the Bell state $|\Psi_-\rangle$ (see [14], Section 2.5) written in terms of eigenstates of $\widehat{\sigma}_z^A$ and $\widehat{\sigma}_z^B$

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|+1\rangle_A \otimes |-1\rangle_B - |-1\rangle_A \otimes |+1\rangle_B) \qquad (79)$$

we find that

$$E(\underset{\rightarrow}{a} \cdot \underset{\rightarrow}{\widehat{\sigma}}^A \otimes \underset{\rightarrow}{b} \cdot \underset{\rightarrow}{\widehat{\sigma}}^B) = -\underset{\rightarrow}{a} \cdot \underset{\rightarrow}{b} \qquad (80)$$

The Bell inequality (75) can be violated for the choice where b_1 and b_2 are orthogonal and a_1, a_2 are parallel to $b_1 + b_2$, $b_1 - b_2$ respectively (see [14], Section 5.1). Furthermore, such a quantum state cannot be described via a hidden variable theory, since Bell inequalities are always satisfied using a hidden variable theory. Experiments have been carried out in optical systems providing

strong evidence for the existence of quantum states that violate Bell inequalities with only a few loopholes remaining(see [30], [15] and [6] for references to experiments). Such violation of Bell inequalities is clearly a *non-classical* feature, since the experiments rule out all local hidden variable theories. As Bell inequalities do not occur for separable states, the experimental observation of a Bell inequality indicates the presence of an entangled state. These violations are not without applications, since such Bell entangled states can be useful in device-independent quantum key distribution [30], [14], [13].

2.8 Non-local Correlations

Another feature of entangled states is that they are associated with *strong correlations* for *observables* associated with *localised sub-systems* that are *well-separated*, a particular example being *EPR correlations* between non-commuting observables. Entangled states can exhibit this particular *non-classical* feature, which again cannot be accounted for via a hidden variable theory.

2.8.1 Local Hidden Variable Theory

Consider two operators $\widehat{\Omega}_A$ and $\widehat{\Omega}_B$ associated with sub-systems A and B. These would be Hermitian if observables are involved, but for generality this is not required. In a local hidden variable theory these would be associated with functions $\Omega_C(\xi)$ ($C = A, B$) of the local hidden variables ξ, with the Hermitean adjoints $\widehat{\Omega}_C^\dagger$ being associated with the complex conjugates $\Omega_C^*(\xi)$. In local hidden variable theory *correlation functions* are given by the following mean values

$$
\langle \Omega_A^* \times \Omega_B \rangle_{LHV} = \int d\xi \, P(\xi) \, \Omega_A^*(\xi) \Omega_B(\xi)
$$

$$
\langle \Omega_A^* \Omega_A \times \Omega_B^* \Omega_B \rangle_{LHV} = \int d\xi \, P(\xi) \, \Omega_A^*(\xi) \Omega_A(\xi) \, \Omega_B^*(\xi) \Omega_B(\xi) \qquad (81)
$$

which then can be shown to satisfy the following *correlation inequality*

$$
|\langle \Omega_A^* \times \Omega_B \rangle_{LHV} |^2 \leq \langle \Omega_A^* \Omega_A \times \Omega_B^* \Omega_B \rangle_{LHV} \qquad (82)
$$

This result is based on the inequality

$$
\int d\xi \, P(\xi) C(\xi) \geq \left(\int d\xi \, P(\xi) \sqrt{C(\xi)} \right)^2 \qquad (83)
$$

for real, positive functions $C(\xi), P(\xi)$ and where $\int d\xi \, P(\xi) = 1$, and which is proved in Appendix 7. In the present case we have $C(\xi) = \Omega_A^*(\xi) \Omega_A(\xi) \, \Omega_B^*(\xi) \Omega_B(\xi)$, which is real, positive. A violation of the inequality in Eq. (82) is an indication of strong correlation between sub-systems A and B.

2.8.2 Non-Entangled State Result

It can be shown that the correlation inequalities are *always* satisfied for non-entangled states. In quantum theory the correlation functions are given by $\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle = Tr(\widehat{\rho}\,\widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B)$ and $\left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle = Tr(\widehat{\rho}\,\widehat{\Omega}_A^\dagger \widehat{\Omega}_B \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B)$. For a non-entangled state of sub-systems A and B we have

$$\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle = \sum_R P_R \left\langle \widehat{\Omega}_A^\dagger \right\rangle_R^A \left\langle \widehat{\Omega}_B \right\rangle_R^B$$

$$\left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle = \sum_R P_R \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \right\rangle_R^A \left\langle \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle_R^B \qquad (84)$$

Now

$$|\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle| \leq \sum_R P_R \,|\left\langle \widehat{\Omega}_A^\dagger \right\rangle_R^A|\,|\left\langle \widehat{\Omega}_B \right\rangle_R^B| \qquad (85)$$

since the modulus of a sum is always less than the sum of the moduli. Using $\left\langle \left(\widehat{\Omega}_C^\dagger - \left\langle \widehat{\Omega}_C^\dagger \right\rangle \right) \left(\widehat{\Omega}_C - \left\langle \widehat{\Omega}_C \right\rangle \right) \right\rangle \geq 0$ with $(C = A, B)$, we obtain the Schwarz inequality - which is true for all states - $\left\langle \widehat{\Omega}_C^\dagger \widehat{\Omega}_C \right\rangle \geq \left\langle \widehat{\Omega}_C^\dagger \right\rangle \left\langle \widehat{\Omega}_C \right\rangle = |\left\langle \widehat{\Omega}_C \right\rangle|^2 = |\left\langle \widehat{\Omega}_C^\dagger \right\rangle|^2$, and hence

$$|\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle| \leq \sum_R P_R \sqrt{\left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \right\rangle_R^A} \sqrt{\left\langle \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle_R^B} \qquad (86)$$

Next we use the inequality

$$\sum_R P_R\, C_R \geq \left(\sum_R P_R \sqrt{C_R} \right)^2 \qquad (87)$$

for real, positive functions C_R, P_R and where $\sum_R P_R = 1$. This inequality, which was used in the paper by Hillery et al [54], is proved in Appendix 7. In the present case we have $C_R = \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \right\rangle_R^A \left\langle \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle_R^B$ so that

$$|\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle|^2 \leq \sum_R P_R \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \right\rangle_R^A \left\langle \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle_R^B = \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle \qquad (88)$$

Thus for a *non-entangled state* we obtain the *correlation inequality*

$$|\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle|^2 = |\left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \right\rangle|^2 \leq \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle \qquad (89)$$

where the general result $\left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle = \left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \right\rangle^*$ has been used. Thus non-entangled states have correlation functions that are consistent with hidden variable theory.

2.8.3 Correlation Violation and Entanglement

Hence if it is found that the correlation inequality is violated $| \left\langle \widehat{\Omega}_A^\dagger \otimes \widehat{\Omega}_B \right\rangle |^2 = | \left\langle \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \right\rangle |^2 > \left\langle \widehat{\Omega}_A^\dagger \widehat{\Omega}_A \otimes \widehat{\Omega}_B^\dagger \widehat{\Omega}_B \right\rangle$ then the state must be entangled, so the correlation inequality violation is also a sufficiency test for entanglement.

3 Identical Particles and Entanglement

We now take into account the situation where systems of *identical particles* are involved. This requires us to give special consideration to the requirement that quantum states in such cases must conform to the *symmetrisation principle* [51]. Further, entanglement is defined as a property that involves systems with two (or more) *sub-systems*, and the definition requires the specification of sub-systems that are *distinguishable* from each other and on which *measurements* can be made. In addition, the sub-systems must be able to exist as separate systems which can in principle be prepared in quantum states for that sub-system alone. This feature is vital to the definition of separable (or non-entangled) states on which the defintion of entangled states is based. These key requirements that the sub-systems must be distinguishable, susceptible to measurements and can exist in separate quantum states are necessary for the concept of entanglement to make physical sense, and will have important consequences for the choice of sub-systems when identical particles are involved. These three key logical requirement for sub-systems rule out considering labelled identical particles as sub-systems and lead to the conclusion that sub-systems must be modes or sets of modes.

3.1 Symmetrisation Principle

Whether *entangled* or *not* the quantum states for systems of *identical particles* must conform to the *symmetrisation principle*, whereby for mixed states the overall density operator has to be invariant under *permutation operators,* or if pure states are involved, the state vector is either unchanged (bosons) or changes sign (fermions) if the permutation operator is odd. Either a first quantisation approach in which the basis states are written as *symmetrised products* of single particle states occupied by *labeled* identical particles can be used, or a second quantisation approach where the basis states are products of *Fock states* for all single particle states (modes), each Fock state specifying the *number* of identical particles occupying the particular mode. In first quantisation the symmetrisation process *removes* any distinction between identical particles, whereas in second quantisation only *mode creation operators* are involved, and these do not involve labeled particles. Symmetrization is built into the definition of the Fock states. The two approaches are equivalent, but as we will see the second quantisation approach is more suited to identifying sub-systems and defining entanglement in systems of identical particles.

It is useful to clarify some of the issues involved by considering a simple example. Since density operators can always be expressed in a diagonal form involving their orthonormal eigenstates $|\Phi\rangle$ with real, positive eigenvalues $P(\Phi)$ as $\widehat{\rho} = \sum_{\Phi} P(\Phi) |\Phi\rangle \langle \Phi|$ and each $|\Phi\rangle$ can always be written as a linear combination of basis vectors $|\Psi\rangle$, we will focus on these basis vectors and their forms in both first and second quantisation. We consider a system with $N = 2$ particles, which may be *identical* and are labeled 1 and 2, or they may be *distinguishable*

and labeled α and β. In each case a particle has a choice of two modes which it may occupy. Thus there are two distinct single particle states (modes) designated as $|A\rangle$ and $|B\rangle$ in the *identical* particle case, and four distinct single particle states (modes) designated as $|A_\alpha\rangle$, $|B_\alpha\rangle$ and $|A_\beta\rangle$, $|B_\beta\rangle$ in the *distinguishable* particle case for particles α and β respectively. The notation in first quantisation is that $|C(i)\rangle$ refers to a vector in which particle i is in mode $|C\rangle$. The notation in second quantisation is that $|n\rangle_C$ refers to a vector where there are n particles in mode $|C\rangle$.

For the case of the *identical* particles we consider *basis states* for two *bosons* or for two *fermions*, which are written in terms of *first quantization* as

$$|\Psi\rangle_{boson} \quad = \quad \frac{1}{\sqrt{2}}(|A(1)\rangle \otimes |B(2)\rangle + |B(1)\rangle \otimes |A(2)\rangle) \tag{90}$$

$$|\Psi\rangle_{fermion} \quad = \quad \frac{1}{\sqrt{2}}(|A(1)\rangle \otimes |B(2)\rangle - |B(1)\rangle \otimes |A(2)\rangle) \tag{91}$$

and clearly satisfy the symmetrization principle. In *second quantization* the basis state in both the fermion and boson cases is

$$|\Psi\rangle_{boson,\,fermion} = |1\rangle_A \otimes |1\rangle_B \tag{92}$$

In both first and second quantisation this basis state involves one identical particle in mode $|A\rangle$ and the other in mode $|B\rangle$.

These examples highlight two possibilities for specifying *sub-systems* for systems of *identical* particles. The two possibilities have *differing* consequences in terms of whether specific pure states are regarded as separable or entangled in terms of the general form in Eq.(1) for separable pure states, depending on whether the first or second quantisation approach is used. The *first option* is to regard the *labeled identical particles* as sub-systems - in which case using first quantisation the boson or fermion basis states in Eqs.(90) and (91) would be regarded as *entangled* states of the two sub-systems consisting of particle 1 and particle 2 [11], [67], [70]. This is a more mathematical approach, and suffers from the feature that the sub-systems are not distinguishable and measurements cannot be made on specifically labelled identical particles. In the case of identical particles the option of regarding labeled identical particles as the sub-systems leads to the concept of *entanglement due to symmetrisation*. In the textbook by Peres ([11], see pp126-128) it is stated that "two particles of the same type are *always* entangled". Peres obviously considers such entanglement is a result of symmetrization. The *second option* would be to regard the *modes* or single particle states as sub-systems [29] - in which case using second quantisation the basis state for both fermions or bosons in Eq.(92) would be regarded as a *separable* state of two sub-systems consisting of modes $|A\rangle$ and $|B\rangle$. This is a more physically based approach, and has the advantage that the sub-systems are distinguishable and measurements can be made on specific modes. Noting that in the example the *same* quantum state is involved with one identical particle in mode $|A\rangle$ and the other in mode $|B\rangle$, the different categorisation is disconcerting. It indicates that a choice must be made in regard

to defining sub-systems when identical particles are involved (see SubSection 3.1.1).

Now consider the case where the particles are *distinguishable*. Each distinguishable particle α, β has its own unique set of modes A_α, B_α, A_β, B_β. There are two cases in which one particle α occupies mode $|A_\alpha\rangle$ or $|B_\alpha\rangle$ and the other particle β occupies mode $|A_\beta\rangle$ or $|B_\beta\rangle$. Basis states analogous to the previous ones are given in *first quantization* as

$$|\Psi\rangle_{dist} = |A_\alpha(\alpha)\rangle \otimes |B_\beta(\beta)\rangle \qquad or \qquad |\Psi\rangle_{dist} = |B_\alpha(\alpha)\rangle \otimes |A_\beta(\beta)\rangle \qquad (93)$$

The somewhat surplus particle labels (α) and (β) have been added for comparison with (90) and (91). The states (95) are not required to satisfy the symmetrization principle since the particles are not identical. Each may be either a boson or a fermion. In *second quantisation* the basis states are

$$\begin{aligned}
|\Psi\rangle_{dist} &= (|1\rangle_{A_\alpha} \otimes |0\rangle_{B_\alpha}) \otimes (|0\rangle_{A_\beta} \otimes |1\rangle_{B_\beta}) \\
&\quad or \\
|\Psi\rangle_{dist} &= (|0\rangle_{A_\alpha} \otimes |1\rangle_{B_\alpha}) \otimes (|1\rangle_{A_\beta} \otimes |0\rangle_{B_\beta})
\end{aligned} \qquad (94)$$

In both first and second quantisation, the first case corresponds to particle α being in mode $|A_\alpha\rangle$ and particle β being in mode $|B_\beta\rangle$ with the other two modes empty, and the second case corresponds to particle α being in mode $|B_\alpha\rangle$ and particle β being in mode $|A_\beta\rangle$ with the other two modes empty.

These examples also highlight two possibilities for specifying sub-systems for systems of *distinguishable* particles. In this case the two possibilities have *similar* consequences in terms of whether specific pure states are regarded as separable or entangled, based on the general form in Eq.(1) for separable pure states, irrespective of whether the first or second quantisation approach is used. Here the *first option* is to regard the *labeled distinguishable particles* as sub-systems - in which case using first quantisation the boson or fermion basis states in Eqs.(93) would be regarded as *separable* states of the two sub-systems consisting of particle α and particle β. The *second option* would be to regard the *modes* or single particle states as sub-systems - in which case using second quantisation the basis state for both fermions or bosons in Eq.(94) would be regarded as a *separable* state of four sub-systems consisting of modes $|A_\alpha\rangle$, $|B_\alpha\rangle$ and $|A_\beta\rangle$, $|B_\beta\rangle$. Both expressions refer to the same quantum state, and the same result regarding separability is obtained in both first and second quantisation, even though the number of sub-systems differ. It indicates that either option may be chosen in regard to defining sub-systems when distinguishable particles are involved. However, it is *simpler* if the same option - particles or modes as sub-systems - is made for treating either identical or distinguishable particle systems and we will adopt this approach.

To highlight the distinction between the identical and distinguishable particles situation, we note that for the two distinguishable particle case treated previously we can also form entangled states from the basis states (93) or (94)

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|A_\alpha(\alpha)\rangle \otimes |B_\beta(\beta)\rangle \pm |B_\alpha(\alpha)\rangle \otimes |A_\beta(\beta)\rangle) \qquad (95)$$

which are similar in mathematical form to (90) and (91) when written in first quantisation, and which are given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}((|1\rangle_{A_\alpha} \otimes |0\rangle_{B_\alpha}) \otimes (|0\rangle_{A_\beta} \otimes |1\rangle_{B_\beta}) \pm (|0\rangle_{A_\alpha} \otimes |1\rangle_{B_\alpha}) \otimes (|1\rangle_{A_\beta} \otimes |0\rangle_{B_\beta}))$$
(96)

when written in second quantisation. However, in this case both the first and second quantisation forms are clearly cases of *entangled* states. Whether they are regarded as entangled states of two sub-systems consisting of particle α and particle β (first option) or entangled states of the four sub-systems consisting of modes $|A_\alpha\rangle$, $|B_\alpha\rangle$ and $|A_\beta\rangle$, $|B_\beta\rangle$ (second option) depends on whether particle or modes are chosen as sub-systems.

Note however that *not* all basis states result in separable/entangled distinctions even in the case of identical particles. For the same two mode, two particle case as considered previously for bosons the basis vectors $|A(1)\rangle \otimes |A(2)\rangle$ or $|B(1)\rangle \otimes |B(2)\rangle$ (first quantisation) or equivalently $|2\rangle_A \otimes |0\rangle_B$ or $|0\rangle_A \otimes |2\rangle_B$ (second quantisation) would be regarded as separable states irrespective of whether particle or modes were chosen as the sub-systems. Entangled states such as $(|A(1)\rangle \otimes |A(2)\rangle \pm |B(1)\rangle \otimes |B(2)\rangle)/\sqrt{2}$ (first quantisation) or equivalently $(|2\rangle_A \otimes |0\rangle_B \pm |0\rangle_A \otimes |2\rangle_B)/\sqrt{2}$ (second quantisation) can also be formed from the two doubly occupied basis states. There are no analogous states for fermions due to the Pauli principle.

It is worth noting that these examples illustrate the general point that just the *mathematical form* of the state vector or the density operator alone is not enough to determine whether a separable or an entangled state is involved. The meaning of the factors involved also has to be taken into account. Failure to realise this may lead to states being regarded as separable when they are not (see SubSection 3.4 for further examples).

In the above discussion the symmetrisation principle was complied with both in the first and second quantisation treatments. It should be noted however that some authors disregard the symmetrisation principle. In describing Bose-Einstein condensates ([67], [94]) consider states of the form

$$\widehat{\rho} = \sum_R P_R \, \widehat{\rho}_R^1 \otimes \widehat{\rho}_R^2 \otimes \widehat{\rho}_R^3 \otimes ...$$
(97)

as defining non-entangled states, where $\widehat{\rho}_R^i$ is a density operator for particle i. However such a state would not in general be allowed, since the symmetrisation principle would be violated unless the $\widehat{\rho}_R^i$ were related. For example, consider the state for two identical bosonic atoms given by

$$\widehat{\rho} = P_{\sigma\xi} \, \widehat{\sigma}^1 \otimes \widehat{\xi}^2 + P_{\theta\eta} \, \widehat{\theta}^1 \otimes \widehat{\eta}^2$$
(98)

and apply the permutation $\widehat{P} = \widehat{P}(1 \leftrightarrow 2)$. The invariance of $\widehat{\rho}$ in general requires $\widehat{\sigma} = \widehat{\xi}$ and $\widehat{\theta} = \widehat{\eta}$, giving $\widehat{\rho} = P_\sigma \, \widehat{\sigma}^1 \otimes \widehat{\sigma}^2 + P_\theta \, \widehat{\theta}^1 \otimes \widehat{\theta}^2$. This is a statistical mixture of two states, one with *both* atoms in state $\widehat{\sigma}$, the other with both atoms

54

in state $\widehat{\theta}$. Thus only special cases of (98) are compatible with the symmetrisation principle. Of course if the atoms were all different (atom 1 a Rb^{87} atom, atom 2 a Na^{23} atom, ..) then the expression (98) would be a valid non-entangled state, but there the atomic sub-systems are distinguishable and symmetrisation is not required. Such authors are really ignoring the symmetrisation principle, and in addition are treating the individual identical particles in the BEC as separate sub-systems - a viewpoint we have described previously and will discuss further in the next SubSubSection. For the present we just point out that valid quantum states must comply with the symmetrisation principle.

3.1.1 Sub-Systems - Particles or Modes ?

As highlighted in the previous SubSection 3.1, when the quantum system involves identical particles the very definition of entanglement itself requires special care in regard to identifying legitimate sub-systems. There is a long-standing debate on the issue, with at present two schools of thought - see reviews such as [30] or [34]. As explained in the previous SubSection, the *first* approach is to identify mathematically labelled individual identical *particles* as the sub-systems [11], [67], [94], [70]. Sub-systems may of course also be sets of such individually labeled particles. This approach leads to the conclusion that *symmetrisation* creates entanglement of identical particles. The *second* approach is to identify single particle states or *modes* that the identical particles may occupy as the sub-systems [29]. The sub-systems may of course also be sets of distinguishable modes. This approach leads to the conclusion that it is *interaction processes* between modes that creates entanglement of distinguishable modes.

The approach based on *particle entanglement* is still being used [70]. As explained in SubSection 3.1 this is not the same as mode entanglement so tests and measures for particle entanglement will differ from those for mode entanglement. A further discussion about the distinction is given in [35]. In a recent paper Killoran et al [95] considered original states such as $(|a0(1)\rangle \otimes |a1(2)\rangle \pm |a0(2)\rangle \otimes |a1(1)\rangle)/\sqrt{2}$ involving two modes $a0$ and $a1$ - which were considered (based on first quantisation) as an *entangled* state for two sub-systems consisting of particles 1 and 2, but would be considered (in second quantisation) as a *separable* state $|1\rangle_{a0} \otimes |1\rangle_{a1}$ for two sub-systems consisting of modes $a0$ and $a1$. In addition there were two modes $b0$ and $b1$ which are intially unoccupied. The particles may be bosons or fermions. They envisaged converting such an input state using *interferometer* processes which couple A modes $a0$ and $a1$ to previously unoccupied B modes $b0$ and $b1$, into an output state - which is different. Projective measurements would then be made on the output state, based on having known numbers of particles in each of the A mode pairs $a0$ and $a1$ and in the B mode pairs $b0$ and $b1$. The projected state with one particle in the A modes and one particle in the B modes would be of the form (in second quantisation) $(|1\rangle_{a0} \otimes |0\rangle_{a1} \otimes |0\rangle_{b0} \otimes |1\rangle_{b1} \pm |0\rangle_{a0} \otimes |1\rangle_{a1} \otimes |1\rangle_{b0} \otimes |0\rangle_{b1})/\sqrt{2}$, which is a bipartite entangled state for the two pairs of modes A and B and is

mathematically of the same form as the first quantisation *form* for the original A modes state considered as an example of *particle* entanglement if the correspondences $|a0(1)\rangle \to |1\rangle_{a0} \otimes |0\rangle_{a1}$, $|a1(2)\rangle \to |0\rangle_{b0} \otimes |1\rangle_{b1}$, $|a0(2)\rangle \to |1\rangle_{b0} |0\rangle_{b1}$ and $|a1(1)\rangle \to |0\rangle_{a0} |1\rangle_{a1}$ are made. Even the minus sign is obtained in the the fermion case. Details are given in Appendix 9.Killoran et al stated that this represented a way of *extracting* the original symmetrization generated entanglement. However, another point of view is that the two mode interferometer process *created* an entangled state from a non-entangled state, and as the final measurements are still based on entanglement of modes it is hard to justify the claim that entanglement due to symmetrization exists as a directly observable basic feature in composite quantum systems - though the mapping identified in [95] is.mathematically correct. Furthermore, all quantum states for identical particles are required to be symmetrized, so if symmetrization causes entanglement it differs from the numerous other controllable processes that produce entanglement by coupling the sub-systems. Since the idea of extracting entanglement due to symmetrisation is of current interest, a fuller discussion of the approach by Killoran et al [95] is set out in Appendix 9.

However, it is generally recognised that sub-systems consisting of *individually labeled* identical particles are *not* amenable to *measurements*. What is distinguishable for systems of identical bosons or fermions is *not* the individual particles themselves - which do not carry labels, boson 1, boson 2, etc. - but the *single particle states* or *modes* that the bosons may occupy. For bosonic or fermionic atoms with several hyperfine components, each component will have its own set of modes. For photons the modes may be specified via wave vectors and polarisations. Although the quantum pure states can be specified via symmetrized products of single particle states occupied by specific particles using a *first quantization* approach, it is more suitable to use *second quantization*. Here, a basis set for the quantum states of such sub-systems are the *Fock states* $|n\rangle_A$ ($n = 0, 1, 2, ...$) etc, which specify the number of identical particles occupying the mode A, etc., so in this approach the mode is the sub-system and the Fock states give different quantum states for this sub-system. Symmetrization is built into the definition of the Fock states, so the symmetrisation principle is *automatically* adhered to. If the atoms were fermions rather than bosons the Pauli exclusion principle would of course restrict $n = 0, 1$ only. In this second quantization approach situations with differing *numbers* of identical particles are recognised as being different *states* of a system consisting of a set of modes, not different *systems* as would be the case in first quantisation. The overall system will be associated with quantum states represented in the theory by density operators and state vectors in *Fock space*, which includes states with total numbers of identical particles ranging from zero in the vacuum state right up to infinity. Finally, the artificial concept of entanglement due to symmetrisation is replaced by the physically realistic concept of entanglement due to mode coupling.

The point of view in which the possible *sub-systems* A, B, etc are *modes* (or *sets* of modes) rather than *particles* has been adopted by several authors ([27], [28], [29]), [13], [44], [45] and will be the approach used here - as in [4]. To empha-

sise - what are or are not entangled in the present treatment involving systems of identical particles are *distinguishable modes* not *labelled-indistinguishable-particles*. Overall, the system is a collection of modes, not particles. Particles are associated with mode *occupancies*, and therefore related to specifying the quantum states of the system, rather than the system itself. In terms of this approach, for non-interacting identical particles at zero temperature, the ground states for Bose-Einstein condensates and Fermi gases trapped in a harmonic potential provide examples of non-entangled states for bosonic and fermionic atoms respectively, when the sub-systems are chosen as the harmonic oscillator (HO) modes. In the bosonic case all the bosons occupy the lowest energy HO state, in the fermionic case one fermion occupies each HO state from the lowest up to a high energy state (the Fermi energy) until all the fermions are accommodated. On the other hand, if one particle position states spatially localised in two different regions are chosen as two sub-systems, then the same zero temperature state for the identical particle system is spatially entangled, as pointed out by Goold et al [33]. Note that in this approach states where there is only a *single atom* may still be entangled states - for example with two spatial modes A, B the states which are a quantum superposition of the atom in each of these modes, such as the Bell state $(|1\rangle_A |0\rangle_B + |0\rangle_A |1\rangle_B)/\sqrt{2}$ are entangled states. For entangled states associated with the EPR paradox or for quantum teleportation, the mode functions may be *localised* in well-separated spatial regions - spooky action at a distance - but spatially overlapping mode functions apply in other situations. This distinction is important in discussions of quantum non-locality. Atoms in states with overall spin zero only have one internal state, but two mode systems can be created for their spatial motion using double-well trap potentials. If the wells are separated then two spatially separated modes can be created for studies of quantum non-locality. On the other hand atoms with spin 1/2 have two internal states, which constitute a two mode system. However these two modes may be associated with the same or overlapping spatial wave functions, in which case studies of quantum non-locality are precuded. These latter situation can however still lead to what is referred to as intrasystem entanglement [80]. Furthermore, as well as being distinguishable the modes can act as *separate systems*, with other modes being ignored. For interacting bosonic atoms this is much harder to accomplish experimentally than for the case of photons, where the relatively slow processes in which photons are destroyed in one EM field mode and created in another may require the presence of atoms as intermediaries. Two bosonic atoms in one mode may collide and rapidly disappear into other modes. However, atomic boson interactions can be made very small via *Feshbach resonance* methods. Near absolute zero the basic physics of a BEC in a single trap potential is describable via a *one mode theory*. Hence with A, B, .. signifying distinct modes, the general non-entangled state is given in Eq. (3) though the present paper mainly involves only two modes.

As pointed out in SubSection 3.1, in the case of systems consisting entirely of *distinguishable* particles the sub-systems may still be regarded as sets of modes, namely those single particle states associated with the particular distinguishable particle. In this case the particle descriptor (He atom, Na atom, ..) is

synonomous with its collection of modes. Here all the sub-system states are one particle states.

3.1.2 Multi-Mode Sub-Systems

As well as the simple case where the sub-systems are all *individual* modes, the concept of entanglement may be *extended* to situations where the sub-systems are *sets of modes*, rather than individual modes, In this case entanglement or non-entanglement will be of these distinct sets of modes. Such a case in considered in SubSection **4.3** of paper II, where *pairs* of modes associated with distinct lattice sites are considered as the sub-systems. Another example is treated in He et al [96], which involves a double well potential with each well associated with two bosonic modes, these pairs of modes being the two sub-systems. Entanglement criteria for the mode pairs based on local spin operators associated with each potential well are considered (sse SubSections **4.2 and 5.3** of paper II). A further example is treated by Heaney et al [97], again involving four modes associated with a double well potential. As in the previous example, each mode pair is associated with the same well in the potential, but here a Bell entanglement test was obtained for pairs of modes in the different wells. The concept of *entanglement of sets of modes* is a straightforward extension of the basic concept of entanglement of individual modes.

3.2 Super-Selection Rule

As well as the symmetrisation principle there is a further requirement that quantum states of systems of identical particles must satisfy - these are known as *super-selection rules*. These rules restrict the allowed quantum states of such systems to those in which the *coherences* between states with differing numbers of particles are zero. This applies at the global level for the overall quantum state, but also - as will be discussed in a later sub-section - to the sub-system states involved in the definition of separable or non-entangled states. The justfication of the SSR at both the global and local level will be considered both in terms of simple physics arguements and in terms of reference frames. Examples of SSR and non-SSR compliant states will be given, both for overall states and for separable states. The validity of the SSR for the case of massive bosons or fermions is generally accepted, but in the case of photons there is doubt regarding their applicability -as will be discussed below. As pointed out in the Introduction, in the case of systems consisting entirely of single *distinguishable* particles the sub-systems may still be regarded as sets of modes, namely those single particle states associated with the particular distinguishable particle. Here all the sub-system states are one particle states and the overall system is an N particle state, so the local and global particle number super-selection rules, though true are irrelevant.

3.2.1 Global Particle Number SSR

The question of what quantum states - entangled or not - are possible in the *non-relativistic quantum physics* of a system of identical *bosonic* particles - such as bosonic *atoms* or *photons* - has been the subject of much discussion. Whether *entangled* or *not* it is generally accepted that there is a *super-selection rule* that prohibits *quantum superposition states* of the form

$$|\Phi\rangle = \sum_{N=0}^{\infty} C_N \, |N\rangle \qquad \widehat{\rho} = \sum_{N=0}^{\infty} |C_N|^2 \, |N\rangle\langle N| + \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} (1-\delta_{N,M})C_N \, C_M^* \, |N\rangle\langle M| \tag{99}$$

being *quantum* states when they involve Fock states $|N\rangle$ with differing total numbers N of particles. The density operator for such a state would involve *coherences* between states with differing N. Although such superpositions - such as the *Glauber coherent state* $|\alpha\rangle$, where $C_N = \exp(-|\alpha|^2/2)\, \alpha^N/\sqrt{N!}$ - do have a useful *mathematical* role, they do *not* represent actual quantum states according to the super-selection rule. The papers by Sanders et al [61] and Cable et al [98] are examples of applying the SSR for optical fields, but also using the mathematical features of coherent states to treat phenomena such as interference between independent lasers. The super-selection rule indicates that the most *general quantum state* for a system of identical bosonic particles can only be of the form

$$\begin{aligned}
\widehat{\rho} &= \sum_{N=0}^{\infty} \sum_{\Phi} P_{\Phi N} \, (|\Phi_N\rangle\langle\Phi_N|) \\
|\Phi_N\rangle &= \sum_i C_i^N \, |N\,i\rangle
\end{aligned} \tag{100}$$

where $|\Phi_N\rangle$ is a quantum superposition of states $|N\,i\rangle$ each of which involves exactly N particles, and where different states with the same N are designated as $|N\,i\rangle$. This state $\widehat{\rho}$ is a statistical mixture of states, each of which contains a specific number of particles. Such a SSR is referred to as a *global* SSR, as it applies to the system as a whole. Mathematically, the global particle number SSR can be expressed as

$$[\widehat{N}, \widehat{\rho}] = 0 \tag{101}$$

where $\widehat{N}$ is the *total number* operator.

3.2.2 Examples of Global Particle Number SSR Compliant States

Examples of a state vector $|\Phi_N\rangle$ for an entangled pure state [28] and a density operator $\widehat{\rho}$ for a non-entangled mixed [33] state for a *two mode bosonic* system, both of which are possible quantum states are

$$|\Phi_N\rangle = \sum_{k=0}^{N} C(N,k) \, |k\rangle_A \otimes |N-k\rangle_B \tag{102}$$

$$\widehat{\rho} = \sum_{k=0}^{N} P(k) \, |k\rangle_A \langle k|_A \otimes |N-k\rangle_B \langle N-k|_B \tag{103}$$

The entangled pure state is a superposition of product states with k bosons in mode A and the remaining $N - k$ bosons in mode B. Every term in the superposition is associated with the same total boson number N. The non-entangled mixed state is a statistical mixture of product states also with k bosons in mode A and the remaining $N - k$ bosons in mode B. Every term in the statistical mixture is associated with the same total boson number N. For the case of a *two mode fermionic* system the Pauli exclusion principle restricts the number of possible fermions to two, with at most one fermion in each mode. Expressions for a state with exactly $N = 2$ fermions are

$$|\Phi_2\rangle = |1\rangle_A \otimes |1\rangle_B \qquad (104)$$
$$\widehat{\rho} = |1\rangle_A \langle 1|_A \otimes |1\rangle_B \langle 1|_B \qquad (105)$$

Neither state is entangled and both are the same pure state since $\widehat{\rho} = |\Phi_2\rangle \langle \Phi_2|$. Although the super-selection rules and symmetrisation principle also applies to fermions, as indicated in the Introduction this paper is focused on bosonic systems, and it will be assumed that the modes are bosonic unless indicated otherwise.

Bell states [14], [13] for $N = 1$ bosons provide important examples of entangled two mode pure quantum states that are compliant with the global particle number SSR. The modes are designated A, B and the Fock states are in general $|n_A, n_B\rangle$. These Bell states may be written

$$\left|\Psi_{AB}^-\right\rangle = \frac{1}{\sqrt{2}}(|0_A, 1_B\rangle - |1_A, 0_B\rangle)$$
$$\left|\Psi_{AB}^+\right\rangle = \frac{1}{\sqrt{2}}(|0_A, 1_B\rangle + |1_A, 0_B\rangle) \qquad (106)$$

Neither of these states is separable. There are also two other two mode Bell states given by

$$\left|\Phi_{AB}^-\right\rangle = \frac{1}{\sqrt{2}}(|0_A, 0_B\rangle - |1_A, 1_B\rangle)$$
$$\left|\Phi_{AB}^+\right\rangle = \frac{1}{\sqrt{2}}(|0_A, 0_B\rangle + |1_A, 1_B\rangle) \qquad (107)$$

These however are not compliant with the global particle number SSR. Linear combinations $(\left|\Phi_{AB}^-\right\rangle + \left|\Phi_{AB}^+\right\rangle)/\sqrt{2} = |0_A, 0_B\rangle$ and $(-\left|\Phi_{AB}^-\right\rangle + \left|\Phi_{AB}^+\right\rangle)/\sqrt{2} = |1_A, 1_B\rangle$ are global particle number SSR compliant and also separable, corresponding to states with $N = 0$ and $N = 2$ bosons respectively.

3.2.3   Super-Selection Rules and Conservation Laws

It is important to realise that such super-selection rules [52] are *different constraints* to those imposed by *conservation laws*, as emphasised by Bartlett et al [60]. For example, the conservation law on total particle number *only* leads to the requirement on the superposition state $|\Phi\rangle$ that the $|C_N|^2$ are time independent, it does *not* require only one C_N being non-zero. They are however related,

as is discussed in Section 3.3.1 and Appendix 10 where the super-selection rules based on particle number are related to invariances of the density operator under changes of phase reference frames. This involves considering groups of phase changing operators $\widehat{T}(\theta_a) = \exp(i\widehat{N}_a\theta_a)$ when considering local particle number SSR for single modes in the context of separable states, or $\widehat{T}(\theta) = \exp(i\widehat{N}\theta)$ when considering global particle number SSR in the context of multimode entangled states. Super-selection rules are broad in their scope, forbidding quantum superpositions of states of systems with differing charge, differing baryon number and differing statistics. Thus a combined system of a hydrogen atom and a helium ion does not exist in quantum states that are linear combinations of hydrogen atom states and helium ion states - the super-selection rules on both charge and baryon number preclude such states. The basis quantum states for such a combined system would involve symmetrised tensor products of hydrogen atom and helium ion states, not linear combinations - symmetrisation being required because the system contains two identical electrons. On the other hand, super-selection rules do not prohibit quantum superpositions of states of systems with differing energy, angular or linear momenta - other physical quantities that may also be conserved. Thus in a hydrogen atom quantum superpositions of states with differing energy and angular momentum quantum numbers are allowed quantum states.

However, conservation laws on total particle number (such as apply in the case of massive bosons) are relevant to showing that multi-mode states generated via total particle number conserving processes from an initial separable state will be global SSR compliant if the sub-systems in the initial state are local particle number SSR compliant, and will not be if the initial state involves a sub-system state that is not local particle number SSR compliant. For simplicity we consider two sub-systems A and B with the initial state

$$\widehat{\rho}(0) = \sum_R P_R \, \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \tag{108}$$

If $\widehat{U}(t)$ is the evolution operator where $\widehat{\rho}(t) = \widehat{U}(t)\widehat{\rho}(0)\widehat{U}(t)^\dagger$ and the processes are number conserving then $[\widehat{N}, \widehat{U}(t)] = 0$. We then have

$$
\begin{aligned}
[\widehat{N}, \widehat{\rho}(t)] &= \widehat{U}(t)\,[\widehat{N}, \widehat{\rho}(0)]\,\widehat{U}(t)^\dagger \\
&= \widehat{U}(t)\sum_R P_R \left([\widehat{N}_A, \widehat{\rho}_R^A] \otimes \widehat{\rho}_R^B + \widehat{\rho}_R^A \otimes [\widehat{N}_B, \widehat{\rho}_R^B]\right)\widehat{U}(t)^\dagger
\end{aligned} \tag{109}
$$

Hence if $\widehat{\rho}_R^A$ and $\widehat{\rho}_R^B$ are local particle number SSR compliant, then $[\widehat{N}_A, \widehat{\rho}_R^A]$ and $[\widehat{N}_B, \widehat{\rho}_R^B]$ are zero, showing that $[\widehat{N}, \widehat{\rho}(t)] = 0$ so the state is global particle number SSR compliant. On the other hand if $[\widehat{N}, \widehat{\rho}(t)] = 0$ we see that $[\widehat{N}, \widehat{\rho}(0)] = \sum_R P_R \left([\widehat{N}_A, \widehat{\rho}_R^A] \otimes \widehat{\rho}_R^B + \widehat{\rho}_R^A \otimes [\widehat{N}_B, \widehat{\rho}_R^B]\right) = 0$. By taking Tr_A and Tr_B of this result gives $\sum_R P_R\,[\widehat{N}_A, \widehat{\rho}_R^A] = \sum_R P_R\,[\widehat{N}_B, \widehat{\rho}_R^B] = 0$. This shows that both of the reduced density operators $\sum_R P_R\widehat{\rho}_R^A$ and $\sum_R P_R\widehat{\rho}_R^B$ must be local particle number SSR compliant, which amounts to requiring the sub-system

density operators to be local particle number SSR compliant. This situation applies even when there is coupling between modes, provided the interaction is number conserving - such as a coupling given by $\widehat{V} = \lambda \widehat{a} \widehat{b}^\dagger + HC$. Analogous results apply for systems of massive bosons if there are more than two modes involved, where again global SSR compliance involves the total particle number since even with interactions there is total number conservation. For example with three modes in coupled BECs, interactions of the form $\widehat{V} = \lambda (\widehat{c})^2 \widehat{a}^\dagger \widehat{b}^\dagger + HC$ in which two bosons are annihilated in mode C and one boson is created in each of modes A and B are consistent with total particle number conservation and lead to global SSR involving the total particle number.

Although outside the focus of this paper, it is worth pointing out that somewhat different considerations apply to photons. Single non-interacting modes, such as are discussed in the context of separable states do have a conservation law for the photon number in that mode. The applicability (or otherwise) of the local particle number SSR for the sub-system density operators in a separable state is discussed in Section 3.2.9. In the case of interacting photonic modes there may be no conservation law associated with total photon number and it may be thought that no global SSR would apply. However, other global SSR involving combinations of the mode photon numbers may still apply. As an example, we consider a three mode situation in a non-degenerate parametric amplifier, where the basic generation process involves one pump photon of frequency $\omega_C = \omega_A + \omega_B$ being destroyed and one photon created in each of modes A and B. The interaction term is $\widehat{V} = \lambda \widehat{c} \widehat{a}^\dagger \widehat{b}^\dagger + HC$. It is straight-forward to show the total *quanta* number operator $\widehat{N}_{tot} = \widehat{N}_A + \widehat{N}_B + 2\widehat{N}_C$ commutes with the Hamiltonian. The situation is analogous to the atom-molecule system treated in Appendix 11. Thus $\widehat{N}_{tot}$ is conserved and we can then consider a group of phase changing operators $\widehat{T}(\theta) = \exp(i\widehat{N}_{tot}\theta)$ and show that there could be a global SSR for the three mode system, but now involving the total quanta number $N_A + N_B + 2N_C$. The pure state which is often used in a quantum treatment of the non-degenerate parametric amplifier $|\Psi\rangle = \sum_n C_n |n\rangle_A \otimes |n\rangle_B \otimes |N - n\rangle_C$ is global SSR compliant in terms of the modified $\widehat{N}_{tot}$, since in every term $N_A + N_B + 2N_C = 2N$ and there are no coherences between terms with different N_{tot}. For the non-degenerate parametric amplifier case an analogous treatment to that for number conserving processes shows that if

$$\widehat{\rho}(0) = \sum_R P_R \, \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes \widehat{\rho}_R^C \tag{110}$$

then using $[\widehat{N}_{tot}, \widehat{U}(t)] = 0$ we have

$$
\begin{aligned}
[\widehat{N}_{tot}, \widehat{\rho}(t)] &= \widehat{U}(t)\,[\widehat{N}_{tot}, \widehat{\rho}(0)]\,\widehat{U}(t)^\dagger \\
&= \widehat{U}(t) \sum_R P_R \left(\begin{array}{c} [\widehat{N}_A, \widehat{\rho}_R^A] \otimes \widehat{\rho}_R^B \otimes \widehat{\rho}_R^C + \widehat{\rho}_R^A \otimes [\widehat{N}_B, \widehat{\rho}_R^B] \otimes \widehat{\rho}_R^C \\ +2\widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes [\widehat{N}_C, \widehat{\rho}_R^C] \end{array} \right) \widehat{U}(t)^\dagger
\end{aligned}
\tag{111}
$$

Hence if $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$ and $\widehat{\rho}_R^C$ are local particle number SSR compliant, then $[\widehat{N}_{tot}, \widehat{\rho}(t)] = 0$ so the state is SSR compliant, but with global total quanta number $\widehat{N}_{tot}$. On the other hand if $[\widehat{N}_{tot}, \widehat{\rho}(t)] = 0$ we find that $\sum_R P_R [\widehat{N}_A, \widehat{\rho}_R^A] = \sum_R P_R [\widehat{N}_B, \widehat{\rho}_R^B] = \sum_R P_R [\widehat{N}_C, \widehat{\rho}_R^C] = 0$. This shows that each of the reduced density operators $\sum_R P_R \widehat{\rho}_R^A$, $\sum_R P_R \widehat{\rho}_R^B$ and $\sum_R P_R \widehat{\rho}_R^C$ must be local particle number SSR compliant, which amounts to requiring the sub-system density operators to be local particle number SSR compliant.

3.2.4 Global SSR Compliant States and Quantum Correlation Functions

We now prove a theorem concerning *quantum correlation functions* for bosonic systems with two modes A and B.

Theorem. If a state is global particle number SSR compliant then all quantum correlation functions $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ for which $n + l \neq m + k$ must be zero.

Proof: If the state is global particle number SSR compliant then if we choose a complete orthonormal set of Fock states $|N, \alpha\rangle$ with $\alpha = 1, 2, ..., d_N$ listing states which are eigenstates of the total number operator $\widehat{N}$ with eigenvalue N we can write the density operator in the form

$$\widehat{\rho} = \sum_N \sum_{\alpha,\beta} P_{\alpha,\beta}^N |N, \alpha\rangle \langle N, \beta| \tag{112}$$

where since $Tr\widehat{\rho} = 1$ we must have

$$1 = \sum_N \sum_\alpha P_{\alpha,\alpha}^N \tag{113}$$

Now $(\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k |N\alpha\rangle$ must be a linear combination of Fock states with N replaced by $N + n + l - m - k$ so we can write

$$(\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k |N, \alpha\rangle = \sum_\gamma C_{\alpha,\gamma}^N(n,m,l,k) |(N + n + l - m - k), \gamma\rangle \tag{114}$$

Hence

$$\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle = Tr \left(\sum_N \sum_{\alpha,\beta} P_{\alpha,\beta}^N \sum_\gamma C_{\alpha,\gamma}^N(n,m,l,k) |(N + n + l - m - k), \gamma\rangle \langle N, \beta| \right) \tag{115}$$

But $Tr(|(N + n + l - m - k), \gamma\rangle \langle N, \beta|) = 0$ unless $n + l - m - k = 0$. Hence

$$\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle = 0 \quad if \ n + l \neq m + k \tag{116}$$

which is the required theorem.

3.2.5 Testing the Super-Selection Rules

The last result for the general two mode quantum correlation function $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ is relevant to the various experimental measurements that are discussed in the accompanying paper II. For example, as we will see $\left\langle \widehat{S}_x \right\rangle_\rho$ is a combination of $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ with $n = 1, m = 0, l = 0, k = 1$ and $n = 0, m = 1, l = 1, k = 0$, and $\left\langle \Delta \widehat{S}_x^2 \right\rangle_\rho = \left\langle \widehat{S}_x^2 \right\rangle_\rho - \left\langle \widehat{S}_x \right\rangle_\rho^2$ would involve terms such as $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ with $n = 2, m = 0, l = 0, k = 2$ and $n = 0, m = 2, l = 2, k = 0$, and $n = 1, m = 1, l = 1, k = 1$ from $\left\langle \widehat{S}_x^2 \right\rangle_\rho$. All of these have $n + l = m + k$, so they can be non-zero for globally SSR compliant states. The question then arises - what sort of quantity of the form $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ could be used to see if the quantum state was not globally SSR compliant? The answer is seen in terms of two corollaries to the last theorem.

Corollary 1. If we find that any of the quantum correlation functions $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ are non-zero when $n + l \neq m + k$ then the state is not global particle number SSR compliant.

This result indicates what type of measurement is needed to see if SSR non compliant states exist. Quantities of the type $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ are measured for which $n + l \neq m + k$. If we find *any* that are non-zero we can then conclude that we have found a state which is *not* global particle number SSR compliant.

Corollary . Measurements of the QCF $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ when $n + l = m + k$ cannot determine whether or not a state is global particle number SSR non-compliant.

If the state is *non-compliant* then its density operator must contain a contribution which allows for *non-zero coherences* between Fock states with different N. We can therefore write the density operator as

$$\widehat{\rho} = \sum_N \sum_{\alpha,\beta} P_{\alpha,\beta}^N \, |N, \, \alpha\rangle \langle N, \, \beta| + \sum_{N \neq M} \sum_{\alpha,\beta} P_{\alpha,\beta}^{N,M} \, |N, \, \alpha\rangle \langle M, \, \beta| \qquad (117)$$

where the second term is the *SSR non-compliant contribution.*

A similar calculation to before for the situation when $n + l = m + k$ gives

$$\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle \;\; = \;\; Tr \left(\sum_N \sum_{\alpha,\beta} P_{\alpha,\beta}^N \sum_\gamma C_{\alpha,\gamma}^N (n, m, l, k) \, |N, \, \gamma\rangle \langle N, \beta| \right)$$

$$+ Tr \left(\sum_{N \neq M} \sum_{\alpha,\beta} P_{\alpha,\beta}^N \sum_\gamma C_{\alpha,\gamma}^N (n, m, l, k) \, |N, \gamma\rangle \langle M, \beta| \right)$$

$$(118)$$

But $Tr(|N,\gamma\rangle\langle M,\beta|) = 0$ for $N \neq M$, so the non-compliant contribution gives zero and as $Tr(|N,\gamma\rangle\langle N,\beta|) = \delta_{\gamma,\beta}$ we end up with

$$\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle = \left(\sum_N \sum_{\alpha,\beta} P^N_{\alpha,\beta} \, C^N_{\alpha,\beta}(n,m,l,k) \right) \qquad (119)$$

which is entirely dependent on the contribution to the density operator that is globally SSR compliant. Measurements of this type with $n + l = m + k$ would *not* respond to the presence of contribution to the density operator that is not globally SSR compliant. The BS measurements discussed in this paper are all of this type, so will not test the super-selection rule.

Hence the *conclusion* is that a quantum correlation function of the form $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m (\widehat{b}^\dagger)^l (\widehat{b})^k \right\rangle$ must be measured for cases where $n + l \neq m + k$ *and* a non-zero measurement result must be found. If it is, then we would have demonstrated that the state is not globally SSR compliant. The simplest case would be to find a non-zero result for $\langle \widehat{a} \rangle_\rho$ or $\langle \widehat{b} \rangle_\rho$.

Similar considerations apply to *local SSR compliance* in the sub-system states. For sub-system a a QCF of the form $\left\langle (\widehat{a}^\dagger)^n (\widehat{a})^m \right\rangle$ must be measured for cases where $n \neq m$ and a non-zero measurement result must be found. If it is, then we would have demonstrated that the state is not locally SSR compliant. The simplest case would be to find a non-zero result for $\langle \widehat{a} \rangle_\rho$.

3.2.6 SSR Justification and No Suitable Phase Reference

There are two types of justification for applying the super-selection rules for systems of identical particles. The first approach is based on simple considerations and will be outlined below in this subsection. The second approach [59], [60], [61], [62], [63], [64], [53], [56], [57], [34] is more sophisticated and involves linking the absence or presence of SSR to whether or not there is a suitable *reference frame* in terms of which the quantum state is described, and is outlined in the next subsection and Appendix 10. The key idea is that SSR are a consequence of considering the description of a quantum state by an external observer (Charlie) whose phase reference frame has an unknown phase difference from that of an observer ((Alice) more closely linked to the system being studied. Thus, whilst Alice's description of the quantum state may violate the SSR, the description of the *same* quantum state by Charlie will not. In the main part of this paper the density operator $\widehat{\rho}$ used to describe the various quantum states will be that of the external observer (Charlie).

3.2.7 SSR Justication and Physics Considerations

A number of *straightforward reasons* have been given in the Introduction for why it is appropriate to apply the superselection rule to exclude quantum superposition states of the form (99) as quantum states for systems of identical particles, and these will now be considered in more detail.

Firstly, no way is known for creating such states. The Hamiltonian for such a system commutes with the total boson number operator, resulting in the $|C_N|^2$ remaining constant, so the quantum superposition state would need to have existed initially. In the simplest case of non-interacting bosonic atoms, the Fock states are also energy eigenstates, such Fock states involve total energies that differ by energies of order the rest mass energy mc^2, so a coherent superposition of states with such widely differing energies would at least seem unlikely in a *non-relativistic theory*, though for massless photons this would not be an issue as the energy differences are of order the photon energy $\hbar\omega$. The more important question is: Is there a non-relativistic quantum process could lead to the creation of such a state? Processes such as the dissociation of M diatomic molecules into up to $2M$ bosonic atoms under Hamiltonian evolution involve entangled atom-molecule states of the form

$$|\Phi\rangle = \sum_{m=0}^{M} C_m \, |M - m\rangle_{mol} \otimes |2m\rangle_{atom} \tag{120}$$

but the reduced density operator for the bosonic atoms is

$$\widehat{\rho}_{atoms} = \sum_{m=0}^{M} |C_m|^2 \, (|2m\rangle \langle 2m|)_{atom} \tag{121}$$

which is a statistical mixture of states with differing atom numbers with no coherence terms between such states. Such statistical mixtures are valid quantum states, corresponding to a lack of a priori knowledge of how many atoms have

been produced. To obtain a quantum superposition state for the atoms *alone*, the atom-molecule state vector would need to evolve at some time into the form

$$|\Phi\rangle = \sum_{m=0}^{M} B_m \, |M-m\rangle_{mol} \otimes \sum_{n=0}^{M} A_{2n} \, |2n\rangle_{atom} \qquad (122)$$

where the separate atomic system is in the required quantum superposition state. However if such a state existed there would be terms with at least one non-zero product of coefficient $B_m A_{2n}$ involving product states $|M-m\rangle_{mol} \otimes |2n\rangle_{atom}$ with $n \neq m$ if the state $|\Phi\rangle$ is not just in the entangled form (120). However, the presence of such a term would mean that the conservation law involving the number of molecules plus two times the number of atoms was violated. This is impossible, so such an evolution is not allowed.

Secondly, no way is known for measuring all the properties of such states, even if they existed. If a state such as (99) did exist then the amplitudes C_N would oscillate with frequencies that differ by frequencies of order $mc^2/\hbar$ (the Compton frequency, which is $\gtrsim 10^{25}$ Hz for massive bosons) even if boson-boson interactions were included To distinguish the phases of the C_N in order to verify the existence of the state, *measurement operators* would need to include terms that also oscillate at relativistic frequencies, and no such measurement operators are known.

Thirdly, there is no need to invoke the existence of such states in order to understand coherence and interference effects..It is sometimes thought that states involving quantum superpositions of number states are needed for discussing *coherence* and *interference properties* of BECs, and some papers describe the state via the Glauber coherent states. However, as Leggett [71] has pointed out (see also Bach et al [99], Dalton and Ghanbari [32]), a highly occupied number state for a single mode with N bosons has coherence properties of high order n, as long as $n \ll N$. The introduction of a Glauber coherent state is *not* required to account for coherence effects. Even the well-known presence of spatial interference patterns produced when two independent BECs are overlapped can be accounted for via treating the BECs as Fock states. The interference pattern is built up as a result of successive boson position measurements [100], [61], [98].

3.2.8   SSR Justification and Galilean Frames ?

Finally, in addition to the previous reasons there is an arguement that has been proposed based on the requirement that the dynamical equations for such non-relativistic quantum systems should be invariant under a *Galilean transformation* which has been proposed [101] as a proof of the super-selection rule for atom number. This approach is linked to the reference frame based justification of SSR (see Appendix 10). However, whilst the paper shows that under a Galilean transformation - corresponding to describing the system from the point of view of an observer moving with a constant velocity **v** with respect to the original observer, and where the two observers have identical clocks - the terms in a superposition state with different numbers N of massive bosons

would oscillate like $\exp i \left(\frac{1}{2} N m \mathbf{v}^2 t\right)/\hbar$, and may be expected if the *same* quantum state is described by a moving observer. This feature alone does not seem to require the super-selection rule, since here the moving observer's reference frame has a well-defined velocity with respect to that attached to the system. However, the moving observer's reference frame may actually have an unknown relative velocity, in which case a twirling operation resulting in the elimination of number state coherences could be involved (see Appendix 10). This will be not be considered further at this stage.

On the other hand, an approach of this kind involving *rotation symmetry* would seem to rule out such states as quantum superpositions of a boson (spin 0) and a fermion.(spin 1/2). Let such a state be prepared in the form $(|F\rangle + |B\rangle)/\sqrt{2}$. Consider an observer whose cartesian reference frame is X, Y, Z. This is a classical system that can be rotated in space. If the observer rotates with his frame through 2π about any axis they are then back in the same position, but the observer now sees the state as $(-|F\rangle + |B\rangle)/\sqrt{2}$. This state is apparently orthogonal to the one observed before the rotation, and this is paradoxical since the observer would be in the same position. Thus there is a super-selection rule excluding states such as $(|F\rangle + |B\rangle)/\sqrt{2}$. A similar argument based on the *time reversal* anti-unitary operator was given by Wick et al [52].

3.2.9   SSR and Photons

Though this paper is focused on massive bosonic atoms the question is whether similar SSR also apply to the optical quantum EM field, which involve *massless* bosons - *photons*. Here the situation is not so clear and we therefore merely present the differing viewpoints in the current literature. Some of the same general reasons for applying the super-selection rule to systems of identical massive bosons also apply here, though the details differ, but others do not. The situation depends also on whether optical or microwave photons are involved. The issue is whether for individual photon modes, states can be prepared that are not local particle number SSR compliant, and if so can the effects of the non-SSR compliant terms be observed and furthermore do we need to invoke the existence of non SSR compliant states to understand interference and coherence effects. As we will see, some SSR non compliant feature needs to be present in order to prepared allegedly non SSR compliant states. Another way of looking at the issue is to ask whether phase reference systems exist for photon modes. In addition, there is the issue for multi-mode situations whether states can be prepared, observed or are needed that are not global particle number SSR compliant. As we will see, SSR may now involve a modified total particle number involving combinations of the mode numbers diferent to the total particle number - because it is these combinations that are conserved in mode interaction processes. In the approach adopted in this paper, the states in question are those described by so-called external observers - not hypothetical observers that are somehow attached to phase reference systems internal to the experiment (see below, Section 3.3). The SSR issue for systems involving massless photons is particularly important in regard to describing entanglement. As explained be-

low in Section 3.4, if separate sub-system states in photonic systems can be prepared with density operators that violate the local particle number SSR, then these so-called "separable but non-local" states [49] would be classified as separable rather than as entangled states. Some of the tests for entanglement described in Paper II for systems of massive boson (such as spin squeezing in any spin component) would then no longer apply for photonic systems, though others (such as the Hillery spin variance test) which do not depend on SSR would still apply.

We first consider the requirement of showing how a non SSR compliant states can be prepared for single modes. In the case of photons, Molmer [102] has argued that the quantum state for a single mode optical laser field operating well above threshold is not a Glauber coherent state, and the density operator would be a statistical mixture of the form (100), with $|\Phi_N\rangle = |N\rangle$ and $P_{\Phi N} = \exp(-\overline{N})\,\overline{N}^N/N!$. Here the density operator is a statistical mixture of photon number states with a Poisson distribution, or equivalently a statistical mixture of equal amplitude coherent states $|\alpha\rangle$ with $\alpha = \sqrt{\overline{N}}\exp(i\phi)$ and all phases ϕ having equal probability. In either form, the quantum state is SSR compliant. In terms of possible processes for preparing states for single mode optical laser fields, this feature is confirmed in theories for single mode lasers involving atomic gain media energised via incoherent pumping processes - there is no well defined optical phase that is imposed on the process. The Scully-Lamb theory (see Mandel and Wolf [103], p935) gives the above threshold steady state density operator for the laser mode can be written in the form of a statistical mixture of number states (somewhat broader than for a Poisson distribution), which again is an SSR compliant state with no well-defined optical phase. Further detailed discussion of laser light generation processes by Pegg and Jeffers [75] confirms this. An alternative approach is presented by Wiseman et al [104], [105], in which the optical laser is treated via a master equation, but where monitoring of the laser environment (difficult!) is required to determine whether certain pure state ensembles - such as those involving coherent states - are physically realisable. The conclusion reached is that for finite self energy the coherent state ensemble is not physically realisable, the closest ensemble being that involving squeezed states, though for zero self energy coherent state ensembles are obtained. On the other hand, microwave photons in single mode high Q cavities can be generated by oscillating electric currents having a well-defined phase. In this case, as shown in experiments on the Jaynes-Cummings model by Rempe et al [106] and Brune et al [107] demonstrating collapses and revivals of Rydberg atom population differences, it is possible to create Glauber coherent states in microwave cavity modes, and the presence of these states are necessary to explain the collapse and revival effects.

We next consider cases where interacting photon modes are involved. The two mode squeezed states generated for example in a non-degenerate parametric amplifier are often written in the form $\sum_n C_n |n\rangle_A |n\rangle_B$, corresponding to the basic generation process in which a pump photon of frequency $\omega_C = \omega_A + \omega_B$ is destroyed and one photon is created in each of modes A and B. Such a state

69

is not even global SSR compliant, but is used in describing various quantum information processes as well as describing two mode squeezing. However, whilst mathematically convenient for treating such applications this do not demonstrate that this two mode pure state has actually been created. This state vector is in fact based on a very simplified version of the process, in which the pump mode is treated classically. If it is treated quantum mechanically and there were N photons initially in mode C, the interaction term $\widehat{V} = \lambda \widetilde{c} \widehat{a}^{\dagger} \widehat{b}^{\dagger} + HC$ would result in a global SSR compliant state vector like $\sum_n C_n |N - n\rangle_C |n\rangle_A |n\rangle_B$, but now involving a total *quanta* number $\widehat{N}_{tot} = \widehat{N}_A + \widehat{N}_B + 2\widehat{N}_C$ (see Section 3.2.3 for details). The state describing modes A and B alone would be $\widehat{\rho}_{AB} = \sum_n |C_n|^2 |n\rangle_A \langle n|_A \otimes |n\rangle_B \langle n|_B$. This state is global SSR compliant in terms of $\widehat{N}_{A,B} = \widehat{N}_A + \widehat{N}_B$ but is not the same as the pure state $\sum_n C_n |n\rangle_A |n\rangle_B$ - which is not SSR compliant. Even more elaborate quantum treatments allowing for irreversible damping processes for all three modes (see for example McNeil et al [108]) that result in non SSR compliant steady state solutions, include assumptions such as the pump mode being coupled to a laser mode that is treated classically - and thus begging the question of whether non SSR compliant states were prepared, since the classical treatment of the laser mode is itself SSR non compliant. We are unaware of any situation where non SSR compliant states are claimed to have been created for *optical photons*, where the theoretical treatment of the preparation process has not assumed the presence of non SSR compliant states for some key sub-system involved - usually in an input pump mode. In Section 3.2.3 we considered a preparation process for the non-degenerate parametric amplifier involving conservation of $\widehat{N}_{tot}$ starting from an initial separable state in which all the sub-system density operators are local particle number SSR compliant and show that this results in a quantum state that is global SSR compliant in terms of total quanta number $\widehat{N}_{tot}$. There must therefore be some non SSR compliant feature in the initial state (which could include pump modes) to produce global non SSR compliant states, so then the issue shifts back to how these non SSR compliant states are prepared in the first place.

Second, there is the requirement of being able to measure the non SSR compliant terms. For the free quantum EM field there is a conservation law for the photon number in each mode, so in a pure state such as in Eq. (99) the $|C_N|^2$ would be time independent. However, for photons the C_N would oscillate with frequencies that only differ by non-relativistic frequencies of order ω rather than the Compton frequency that applies for massive bosons, so the arguement against being able to detect coherent states based on this frequency being so large that the oscillations cannot be followed do not necessarily apply. Clocks that can follow microwave oscillations are common-place, and even at optical frequencies the development of atomic clocks based on optical atomic transitions that may enable optical frequency oscilations to be observed. So this consideration does not rule out non SSR compliant states, even for optical photons.

Finally, we consider the requirement of non SSR compliant states being needed to explain interference, coherence effects etc. We need to distinguish the sitation where it is mathematically convenient to invoke SSR non compliant states to explain these effects from the situation where it is essential to do so. Thus it may be convenient to explain the presence of interference patterns in position measurements for bosons from two independent BECs by choosing Glauber coherent states to represent their states, but as pointed out in Section 3.2.7 such interference patterns are accounted in terms of Fock states, together with quantum interference of probability amplitudes associated with bosons being taken from the two different sources (see Refs. [100], [61], [98]). In fact, the more detailed feature that although the separation of the peaks is well defined the actual position of the peaks are random, is inconsistent with the Glauber coherent state description. We also point out below (see Section 3.3.2 and Appendix 11) that the interpretation of Ramsey fringes in a proposed experiment to detect a coherent superposition of an atom and a molecule does not show that such a state was created or that the BEC involved had to be described by a Glauber coherent state. Many experiments in which coherence, interference effects are observed do not depend on SSR non compliant states being created. Optical interference and coherence effects can also be explained without invoking Glauber coherent states, as shown by Molmer [102] and in other papers such as [61]. In regard to two mode squeezing in the non-degenerate parametric amplifier described above, the observation of squeezing effects is often discussed in terms of SSR non compliant states of the form $\sum_n C_n \left|n\right\rangle_A \left|n\right\rangle_B$

(see Ref. [15] for example). However, as may be seen from the experimental paper of Ou et al [109], the way in which two mode squeezing is observed in the non-degenerate parametric amplifier involves generating the pump field by frequency doubling from a lower frequency laser. That lower frequency laser is also used to provide the local oscillator fields for the homodyne measurements on modes A and B used to detect squeezing - these modes are coupled to the local oscillator fields using beam splitters. The original lower frequency laser acts as an internal phase reference for the overall experiment, as may be seen in Fig 2 of Ref [109] which involves the relative phase between the local oscillator and the squeezed input field. But as there is no external phase reference system involved, only the relative phases of the A, B and local oscillator modes are well defined, and not the overall phase as would be required for preparing non SSR compliant states. Again the convenient use of SSR non compliant states to understand experiments in which only internal phase references are involved does not show that SSR non compliant states are necessary to interpret the experiments. A similar arrangement occurs for the degenerate parametric amplifier experiment of Wu et al [110], where Fig 2 clearly shows how the local oscillator field derives from the original Nd-YAG laser.

Perhaps the best way to approach the question of whether SSR compliance is required for optical photons for example (see next sub-section and SubSection 10.4 in Appendix 10) involves the consideration of phase reference frames. The quantum state of a single mode laser may be described as a Glauber coherent

state by an observer (Alice) with one reference frame, but would be described as a statistical mixture of photon number states by another observer (Charlie) with a different reference frame whose phase reference is completely unrelated to the previous one. This arguement against the presence of coherent state in Charlie's viewpoint is only overcome if inter-related phase reference frames at the relevant photon frequencies actually exist.

3.3 Reference Frames and Violations of Superselection Rules

Challenges to the requirement for quantum states to be consistent with super-selection rules have occured since the 1960's when Aharonov and Susskind [59] suggested that coherent superpositions of different charge eigenstates could be created. It is argued that super-selection rules are not a fundamental requirement of quantum theory, but the restrictions involved could be lifted if there is a suitable system that acts as a *reference* for the coherences involved - [59], [60], [61], [62], [63], [64], [53], [56], [57], [34] provide discussions regarding reference systems and SSR.

3.3.1 Linking SSR and Reference Frames

The discussion of the super-selection rule issue in terms of reference systems is quite complex and too lengthy to be covered in the body of this paper. However, in view of the wide use of the reference frame approach a full outline is presented in Appendix 10. The key idea is that there are two observers - Alice and Charlie - who are describing the same prepared system in terms of their own reference frames and hence their descriptions involve two different quantum states. The reference systems are *macroscopic systems* in states where the behaviour is essentially *classical*, such as large magnets that can be used to define *cartesian axes* or BEC in Glauber coherent states that are introduced to define a *phase reference*. The relationship between the two reference systems is represented by a *group* of *unitary transformation operators* listed as $\widehat{T}(g)$, where the particular transformation (translation or rotation of cartesian axes, phase change of phase references, ..) that changes Alice's reference system into Charlie's is denoted by g. Alice describes the quantum state via her density operator, whereas Charlie is the *external* observer whose specification of the *same* quantum state via his density operator is of most interest. There are two cases of importance, *Situation A* - where the relationship between Alice's and Charlie's reference frame is is *known* and specified by a *single* parameter g, and *Situation B* - where on the other hand the relationship between frames is completely *unknown*, all possible transformations g must be given equal weight. Situation A is not associated with SSR, whereas Situation B leads to SSR. The relationship between Alice's and Charlie's density operators is given in terms of the transformation operators (see Eq. (225) for Situation A and Eq. (226) for Situation B). In Situation B there is often a qualitative change between Alice's and Charlie's description of the same quantum state, with pure states as described by Alice becoming mixed states when described by Charlie. It is

Situation B with the *U(1)* transformation group - for which *number operators* are the *generators* - that is of interest for the *single* or *multi-mode* systems involving *identical bosons* on which the present paper focuses. An example of the qualitative change of behaviour for the single mode case is that *if* it is *assumed* that Alice could prepare the system in a Glauber coherent pure state - which involves SSR breaking coherences between differing number states - then Charlie would describe the same state as a Poisson statistical mixture of number states - which is consistent with the operation of the SSR. Thus the SSR applies in terms of external observer Charlie's description of the state. This is how the dispute on whether the state for single mode laser is a coherent state or a statistical mixture is resolved - the two descriptions apply to different observers - Alice and Charlie. On the other hand there are quantum states such as Fock states and Bell states which are described the same way by both Alice and Charlie, even in Situation B. The general justification of the SSR for Charlie's density operator description of the quantum state in Situation B is derived in terms of the *irreducible representations* of the transformation group, there being no coherences between states associated with differing irreducible representations (see Eq. (250)). For the particular case of the *U(1)* transformation group the irreducible representations are associated with the total *boson number* for the system or sub-system, hence the SSR that prohibits coherences between states where this number differs. Finally, it is seen that if Alice describes a general non-entangled state of sub-systems - which being separable have their own reference frames - then Charlie will also describe the state as a non-entangled state and with the same probability for each product state (see Eqs. (255) and (256)). For systems involving *identical bosons* Charlie's description of the sub-system density operators will only involve density operators that conform to the SSR. This is in accord with the key idea of the present paper.

3.3.2   Can Coherent Superpositions of Atom and Molecules Occur ?

Based around the reference frame approach Dowling et al [111] and Terra Cunha et al [29] propose processes using a BEC as a reference system that would create a coherent superposition of an atom and a molecule, or a boson and a fermion [111]. Dunningham et al [112] consider a scheme for observing a superposition of a one boson state and the vacuum state. Obviously if super-selection rules can be overcome in these instances, it might be possible to *produce* coherent superpositions of Fock states with differing particle numbers such as Glauber coherent states, though states with $\overline{N} \sim 10^8$ would presumably be difficult to produce. However, detailed considerations of such papers indicate that the states actually produced in terms of Charlie's description are statistical mixtures consistent with the super-selection rules rather than coherent superpositions, which are only present in Alice's description of the state (see Appendix 10). Also, although coherence and interference effects are demonstrated, these can also be accounted for without invoking the presence of coherent superpositions that violate the super-selection rule. As the paper by Dowling et al [111] entitled "Observing a coherent superposition of an atom and a molecule."

is a good example of where the super-selection rules are challenged, the key points are described in Appendix 11. Essentially the process involves one atom A interacting with a BEC of different atoms B leading to the creation of one molecule AB, with the BEC being depleted by one B atom. There are three stages in the process, the first being with the interaction that turns separate atoms A and B into the molecule AB turned on at Feshbach resonance for a time t related to the interaction strength and the mean number of bosons in the BEC reference system, the second being free evolution at large Feshbach detuning Δ for a time τ leading to a phase factor $\phi = \Delta\,\tau$, the third being again with the interaction turned on at Feshbach resonance for a further time t. However, it is pointed out in Appendix 11 that Charlie's description of the state produced for the atom plus molecule system is merely a statistical mixture of a state with one atom and no molecules and a state with no atom and one molecule, the mixture coefficients depending on the phase ϕ imparted during the process. However a coherent superposition is seen in Alice's description of the final state, though this is not surprising since a SSR violating initial state was assumed. The feature that in Charlie's description of the final state no coherent superposition of an atom and a molecule is produced in the process is not really surprising, because of the averaging over phase differences in going from Alice's reference frame to Charlie's. It is the dependence on the phase ϕ imparted during the process that demonstates coherence (Ramsey interferometry) effects, but it is shown in Appendix 11 that exactly the same results can be obtained via a treatment in which states which are coherent superpositions of an atom and a molecules are never present, the initial BEC state being chosen as a Fock state. In terms of the description by an external observer (Charlie) the claim of violating the super-selection rule has not been demonstrated via this particular process.

3.3.3   Detection of SSR Violating States

Whether such super-selection rule violating states can be *detected* has also not been justified. For example, consider the state given by a superposition of a one boson state and the vacuum state (as discussed in [112]). We consider an interferometric process in which one mode A for a two mode BEC interferometer is initially in the state $\alpha\,|0\rangle + \beta\,|1\rangle$, and the other mode B is initially in the state $|0\rangle$ - thus $|\Psi(i)\rangle = (\alpha\,|0\rangle + \beta\,|1\rangle)_A \otimes |0\rangle_B$ in the usual occupancy number notation, where $|\alpha|^2 + |\beta|^2 = 1$. The modes are first coupled by a beam splitter, then a free evolution stage occurs for time τ associated with a phase difference $\phi = \Delta\tau$ (where $\Delta = \omega_B - \omega_A$ is the mode frquency difference), the modes are then coupled again by the beam splitter and the probability of an atom being found in modes A, B finally being measured. The probabilities of finding one atom in modes A, B respectively are found to only depend on $|\beta|^2$ and ϕ. Details are given in Appendix 11. There is no dependence on the relative phase between α and β, as would be required if the superposition state $\alpha\,|0\rangle + \beta\,|1\rangle$ is to be specified. Exactly the same detection probabilities are obtained if the initial state is the mixed state $\widehat{\rho}(i) = |\alpha|^2 (|0\rangle_A \langle 0|_A \otimes |0\rangle_B \langle 0|_B) + |\beta|^2 (|1\rangle_A \langle 1|_A \otimes$

$|0\rangle_B \langle 0|_B$), in which the vacuum state for mode A occurs with a probability $|\alpha|^2$ and the one boson state for mode A occurs with a probability $|\beta|^2$. In this example the proposed coherent superposition associated with the super-selection rule violating state would not be detected in this interferometric process, nor in the more elaborate scheme discussed in [112].

Of course, the claim that in isolated systems of *massive* particles it is not possible in *non-relativistic quantum physics* to create states that violate the particle number SSR - either for the sub-system states in a separable state or for any quantum state of the overall system - can be questioned. Ideally the claim should be *tested* by experiment, in particular when the number of particles is large in view of the interest in *macroscopic entanglement* since the Schrodinger cat was first described. The simplest situation would be to test whether states that violate the (local) particle number SSR could be created for a *single mode system*. Clearly, a specific *proposal* for an experiment in which the SSR *could* be violated is required, but to our knowledge no such proposal has been presented. Bose-Einstein condensates, in which all the bosons can occupy a single mode would seem an ideal candidate as a suitable bosonic system, and the *Glauber coherent state* is an example of a non SSR compliant state. For fermions, the Pauli exclusion principle would limit the number of fermions in a one mode system to be zero or one, but coherent superpositions of a zero and one fermion state are examples of non-SSR compliant pure states. As pointed out above, some authors such as [54], [55], [49], [50], base their definition of entanglement by allowing for the possible presence of non-SSR compliant sub-system states when defining separable states. The approach in Refs. [4] is based on the *physical assumption* that states that are non-compliant with particle number SSR - both local and global - do *not* come into the realm of non-relativistic quantum physics, in which the concept of entanglement is useful. Until *clear evidence* is presented that non-SSR- compliant states *can* be prepared, and in view of the theoretical reasons why they *cannot*, it seems preferable to base the *theory of entanglement* on their *absence* when defining separable and entangled quantum states.

3.4   Super-Selection Rule - Separate Sub-Systems

In this sub-section the important case of SSR in *separable* states will be dealt with, since this is key to understanding what entangled states are allowed in systems involving identical particles. This forms the basis for the treatment of entanglement tests presented in the second part of this review (Paper II).

3.4.1   Local Particle Number SSR

We now consider the role of the super-selection rule for the case of *non-entangled* states. The global super-selection rule on *total particle number* has restricted the physical quantum state for a system of identical bosons to be of the form (100). Such states may or may not be entangled states of the modes involved. The question is - do similar restrictions involving the *sub-system particle number*

apply to the modes, considered as *separate* sub-systems in the definition of non-entangled states ? The viewpoint in this paper is that this is so. Note that applying the SSR on the separate sub-system density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... is *only* in the context of non-entangled states. Such a SSR is referred to as a *local* SSR, as it applies to each of the separate sub-systems. Mathematically, the local particle number SSR can be expressed as

$$[\widehat{N}_X, \widehat{\rho}_R^X] = 0 \tag{123}$$

where $\widehat{N}_X$ is the *number* operator for sub-system $X = A, B, ...$ The SSR restriction is based on the proposition that the density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... for the separate sub-systems A, B, ... should themselves represent possible *quantum states* for each of the sub-systems, considered as a *separate system* and thus be required to satisfy the super-selection rule that forbids quantum superpositions of Fock states with differing boson numbers. Note that if the local particle number SSR applies in each sub-system the global particle number SSR applies to any separable state. The proof is trivial and just requires showing that $[\widehat{N}, \widehat{\rho}_{sep}] = 0$

The justification of applying the *local particle number super-selection rule* to the density operators $\widehat{\rho}_R^a$, $\widehat{\rho}_R^b$, ... for the sub-system quantum states that occur in any separable state is simply that these are possible quantum states of the sub-systems when the latter are considered as separate quantum systems before being combined as in the Werner protocol [5] to form the separable state. Hence all the justifications based either on simple physical considerations or phase reference systems that were previously invoked for the density operator $\widehat{\rho}$ of any general quantum states of the combined sub-systems to establish the *global particle number super-selection rule* apply equally well here. No more need be said. It is contended that expressions for the non-entangled quantum state $\widehat{\rho}$ in which $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, $\widehat{\rho}_R^C$... were *not* allowed quantum states for the sub-systems would only be of mathematical interest.

Applying the local particle number SSR to the sub-system density operators for non-entangled states is discussed in papers by Bartlett et al [47], [53] as one of several *operational approaches* for defining entangled states. As pointed out above other authors [49], [50] define separable (and hence entangled) states differently by specifically allowing sub-system density operators that are *not* consistent with the local particle number SSR, though the overall density operator is globally SSR consistent. The corresponding overall states are termed *separable but non-local*, and states that they would regard as separable would here be regarded as entangled. Examples of such states are given in Eqs. (130) and (132). There are also other authors [54], [55] who define separable (and hence entangled) states via (3) but leave unspecified whether the sub-system density operators are consistent or inconsistent with the local particle number SSR. Note that any inequalities involving measured quantities that are found for separable states in which local SSR compliance is *neglected* must also apply to separable states where it is *required*. The consequent implications for entanglement tests where local particle number SSR compliance is required is

discussed in the accompanying paper II. Hence, in this paper we are advocating a *revision* to a *widely held notion* of entanglement in identical particle systems, the consequence being that the set of entangled states is now much *larger*. This is a *key idea* in this paper - not only should super-selection rules on particle numbers be applied to the the *overall* quantum state, entangled or not, but it *also* should be applied to the density operators that describe states of the modal *sub-systems* involved in the general definition of *non-entangled* states. The reasons for adopting this viewpoint have been discussed above - basically it is because in separable states the sub-system density operators must represent possible quantum states for the sub-systems considered as isolate quantum systems, so the general reasons for applying the SSR will apply to these density operators also. Apart from the papers by Bartlett et al [47], [53] we are not aware that this definition of non-entangled states has been invoked previously, indeed the opposite approach has been proposed [49], [50]. However, the idea of considering whether sub-system states should satisfy the local particle number SSR has been presented in several papers - [49], [50], [47], [53], [56], [57], [58], mainly in the context of pure states for bosonic systems, though in these papers the focus is on issues other than the definition of entanglement, such as quantum communication protocols [49], multicopy distillation [47], mechanical work and accessible entanglement [56], [57] and Bell inequality violation [58]. However, there are a number of papers that do not apply the SSR to the sub-system density operators, and those that do have not studied the consequences for various entanglement tests. These tests are also discussed in the accompanying paper II.

3.4.2 Criterion for Local and Global SSR in Separable States

Theorem. A necessary and sufficient condition for all separable states for a given set of sub-system density operators $\widehat{\rho}_R^a, \widehat{\rho}_R^b$ to be global particle number SSR compliant is that all such sub-system states are local particle number SSR compliant.

We first note that

$$[\widehat{N}, \widehat{\rho}] = \sum_R P_R([\widehat{n}_a, \widehat{\rho}_R^a] \otimes \widehat{\rho}_R^b + \widehat{\rho}_R^a \otimes [\widehat{n}_b, \widehat{\rho}_R^b]) \tag{124}$$

Necessity: If the state $\widehat{\rho}$ is globally SSR compliant then $[\widehat{N}, \widehat{\rho}] = 0$. Taking the trace of both sides of (124) over sub-system space b, using $Tr_b(\widehat{\rho}_R^b) = 1$ and $Tr_b([\widehat{n}_b, \widehat{\rho}_R^b]) = 0$ and then repeating the process for sub-system space a gives

$$0 = \sum_R P_R([\widehat{n}_a, \widehat{\rho}_R^a]) \qquad 0 = \sum_R P_R([\widehat{n}_b, \widehat{\rho}_R^b]) \tag{125}$$

which are operator equation in sub-system spaces a and b respectively.

The P_R are not independent, satisfying $\sum_R P_R = 1$. By choosing a particular

P_S we can write the last equation for sub-system a as

$$0 = \sum_{R \neq S} P_R([\widehat{n}_a, \widehat{\rho}_R^a]) + (1 - \sum_{R \neq S} P_R)([\widehat{n}_a, \widehat{\rho}_S^a]) \tag{126}$$

where the remaining P_R are now independent. Differentiating the last equation with respect to P_R then gives

$$0 = [\widehat{n}_a, \widehat{\rho}_R^a] - [\widehat{n}_a, \widehat{\rho}_S^a] \tag{127}$$

for any two different R and S. Thus all the $[\widehat{n}_a, \widehat{\rho}_S^a]$ must be the same. Using $0 = \sum_R P_R([\widehat{n}_a, \widehat{\rho}_R^a])$ again with equal $[\widehat{n}_a, \widehat{\rho}_R^a]$ and $\sum_R P_R = 1$ we then see that all $[\widehat{n}_a, \widehat{\rho}_R^a]$ must be zero. Similar considerations show that $[\widehat{n}_b, \widehat{\rho}_R^b] = 0$.

As these results apply for any choice of the P_R and of the $\widehat{\rho}_R^a$, $\widehat{\rho}_R^b$ we can then conclude that

$$[\widehat{n}_a, \widehat{\rho}_R^a] = 0 \qquad [\widehat{n}_b, \widehat{\rho}_R^b] \tag{128}$$

which establishes that the sub-system states are local particle number SSR compliant.

Note that the proof depended on the choice of the P_R being arbitrary apart from $\sum_R P_R = 1$. If the P_R are fixed then although we can show that $0 = \sum_R P_R([\widehat{n}_a, \widehat{\rho}_R^a]) = \sum_R P_R([\widehat{n}_b, \widehat{\rho}_R^b])$, the steps leading to $[\widehat{n}_a, \widehat{\rho}_R^a] = [\widehat{n}_b, \widehat{\rho}_R^b] = 0$ do not follow. The four sub-system states in Section 3.4.3 where all $P_R = 1/4$ are not local particle number SSR compliant even though the overall state is global particle number SSR compliant. This would not be the case if any of the P_R differed from $1/4$.

Sufficiency: If the sub-system states are local particle number SSR compliant then $[\widehat{n}_a, \widehat{\rho}_R^a] = [\widehat{n}_b, \widehat{\rho}_R^b] = 0$. It then follows from (124) that

$$[\widehat{N}, \widehat{\rho}] = 0 \tag{129}$$

which establishes that the separable state is global particle number SSR compliant. This conclusion applies for arbitrary P_R.

3.4.3 Global but not Local Particle Number SSR Compliant States

However, it should be noted that some authors [49], [50] consider sub-system density operators in the context of two mode systems which comply with the global particle number SSR but not the local particle number SSR. Such a case involving four *zero and one boson superpositions* is presented by Verstraete et al [49], [50]. The overall density operator is a statistical mixture

$$\begin{aligned}
\widehat{\rho} &= \frac{1}{4}(|\psi_1\rangle \langle\psi_1|)_A \otimes |\psi_1\rangle \langle\psi_1|)_B + \frac{1}{4}(|\psi_i\rangle \langle\psi_i|)_A \otimes |\psi_i\rangle \langle\psi_i|)_B \\
&\quad + \frac{1}{4}(|\psi_{-1}\rangle \langle\psi_{-1}|)_A \otimes |\psi_{-1}\rangle \langle\psi_{-1}|)_B + \frac{1}{4}(|\psi_{-i}\rangle \langle\psi_{-i}|)_A \otimes |\psi_{-i}\rangle \langle\psi_{-i}|)_B
\end{aligned} \tag{130}$$

where $|\psi_{\omega}\rangle = (|0\rangle + \omega|1\rangle)/\sqrt{2}$, with $\omega = 1, i, -1, -i$. The $|\psi_{\omega}\rangle$ are superpositions of zero and one boson states and consequently the local particle number SSR is violated by each of the sub-system density operators $|\psi_{\omega}\rangle\langle\psi_{\omega}|)_A$ and $|\psi_{\omega}\rangle\langle\psi_{\omega}|)_B$. Although the expression in Eq.(130) is of the form in Eq.(3), the subsystem density operators $|\psi_{\omega}\rangle\langle\psi_{\omega}|)_A$ and $|\psi_{\omega}\rangle\langle\psi_{\omega}|)_B$ do not comply with the local particle number SSR, so in the present paper and in [4] the state would be regarded as *entangled*. However, Verstraete et al [49], [50] regard it as separable. They refer to such a state as *separable but nonlocal*.

On the other hand, the global particle number SSR is obeyed since the density operator can also be wriiten as

$$\begin{aligned}
\widehat{\rho} &= \frac{1}{4}(|0\rangle\langle0|)_A \otimes |0\rangle\langle0|)_B + \frac{1}{4}(|1\rangle\langle1|)_A \otimes |1\rangle\langle1|)_B \\
&+ \frac{1}{2}(|\Psi_+\rangle\langle\Psi_+|)_{AB}
\end{aligned} \tag{131}$$

where $|\Psi_+\rangle_{AB} = (|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B)/\sqrt{2}$. This is a statistical mixture of $N = 0, 1, 2$ boson states. Note that Eq.(131) indicates that the state could be prepared as a mixed state containing two terms that comply with the local particle number SSR in each of the sub-systems plus a term which is an entangled state of the two sub-systems. The presence of an entangled state in such an obvious preparation process challenges the description of the state as being separable.

To further illustrate some of the points made about super-selection rules - local and global - it is useful to consider a second specific case also presented by Verstraete et al [49], [50]. This *mixture* of *two mode coherent states* is represented by the two mode density operator

$$\begin{aligned}
\widehat{\rho} &= \int \frac{d\theta}{2\pi} |\alpha, \alpha\rangle\langle\alpha, \alpha| \\
&= \int \frac{d\theta}{2\pi} (|\alpha\rangle\langle\alpha|)_A \otimes (|\alpha\rangle\langle\alpha|)_B
\end{aligned} \tag{132}$$

where $|\alpha\rangle_C$ is a one mode coherent state for mode $C = A, B$ with $\alpha = |\alpha|\exp(-i\theta)$, and modes A, B are associated with bosonic annihilation operators $\widehat{a}, \widehat{b}$. The magnitude $|\alpha|$ is fixed.

This density operator *appears* to be that for a non-entangled state of modes A, B in the form
$$\widehat{\rho} = \sum_R P_R \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \tag{133}$$

with $\sum_R P_R \to \int \frac{d\theta}{2\pi}$ and $\widehat{\rho}_R^A \to (|\alpha\rangle\langle\alpha|)_A$ and $\widehat{\rho}_R^B \to (|\alpha\rangle\langle\alpha|)_B$. However although this choice of $\widehat{\rho}_R^A, \widehat{\rho}_R^B$ satisfy the Hermitiancy, unit trace, positivity features they do *not* conform to the requirement of satisfying the (*local*) sub-system

boson number *super-selection rule*. From Eq. (132) we have

$$\langle n|\left(|\alpha\rangle\langle\alpha|\right)|m\rangle_A = \exp(-|\alpha|^2)\frac{\alpha^n}{\sqrt{n!}}\frac{(\alpha)^{*m}}{\sqrt{m!}}$$

$$\langle p|\left(|\alpha\rangle\langle\alpha|\right)|q\rangle_B = \exp(-|\alpha|^2)\frac{\alpha^p}{\sqrt{p!}}\frac{(\alpha)^{*q}}{\sqrt{q!}} \tag{134}$$

so clearly for each of the separate modes there are *coherences* between Fock states with differing boson occupation numbers. In the approach in the present paper the density operator in Eq. (132) does *not* represent a non-entangled state. However, in the papers of Verstraete et al [49], [50], Hillery et al [54], [55] and others it would represent an allowable non-entangled (separable) state. Indeed, Verstraete et al [49] specifically state ".., this state is *obviously* separable, though the states $|\alpha\rangle$ are incompatible with the (local) super-selection rule.". Verstraete et al [49] introduce the state defined in Eq. (132) as an example of a state that is separable (in their terms) but which cannot be prepared locally, because it is incompatible with the local particle number super-selection rule.

The *mixture* of *two mode coherent states* does of course satisfy the *total* or *global* boson number super-selection rule. The matrix elements between two mode Fock states are

$$(\langle n|_A \otimes \langle p|_B)\,\widehat{\rho}\,(|m\rangle_A \otimes |q\rangle_B) = \exp(-2|\alpha|^2)\frac{|\alpha|^{n+m}}{\sqrt{n!}\sqrt{m!}}\frac{|\alpha|^{p+q}}{\sqrt{p!}\sqrt{q!}}\int\frac{d\theta}{2\pi}\,\exp(-i(n-m+p-q)\theta)$$

$$= \exp(-2|\alpha|^2)\frac{|\alpha|^{n+m}}{\sqrt{n!}\sqrt{m!}}\frac{|\alpha|^{p+q}}{\sqrt{p!}\sqrt{q!}}\,\delta_{n+p,m+q} \tag{135}$$

These overall matrix elements are zero unless $n+p = m+q$, showing that there are *no coherences* between two mode Fock states where the total boson number differs. The mixture of two mode coherent states has the interesting feature of providing an example of a two mode state which satisfies the global but not the local super-selection rule.

The *reduced density operators* for modes A, B are

$$\widehat{\rho}_A = \int\frac{d\theta}{2\pi}\,(|\alpha\rangle\langle\alpha|)_A \qquad \widehat{\rho}_B = \int\frac{d\theta}{2\pi}\,(|\alpha\rangle\langle\alpha|)_B$$

and a straightforward calculation gives

$$\widehat{\rho}_A = \exp(-|\alpha|^2)\sum_n\frac{|\alpha|^{2n}}{n!}\,(|n\rangle\langle n|)_A \qquad \widehat{\rho}_B = \exp(-|\alpha|^2)\sum_p\frac{|\alpha|^{2p}}{p!}\,(|p\rangle\langle p|)_B$$

which are statistical mixtures of Fock states with the expected Poisson distribution associated with coherent states. This shows that the reduced density operators *are* consistent with the separate mode local super-selection rule, whereas the density operators $\widehat{\rho}_R^A = (|\alpha\rangle\langle\alpha|)_A$, $\widehat{\rho}_R^B = (|\alpha\rangle\langle\alpha|)_B$ are *not* . Later we will revisit this example in the context of entanglement tests.

Note that if a twirling operation (see Eq.(267)) were to be applied to mode A, the result would be equivalent to applying two independent twirling operations

to each mode. In this case the density operator for each mode is a Poisson statistical mixture of number states, so each mode has a density operator that complies with the local particle number SSR.

3.4.4 Particle Entanglement Measure

Wiseman et al have also treated entanglement for pure states [39] and mixed states [111] in identical particle systems, applying both the symmetrization principle and super-selection rules, invoking the arguement that *without* a phase reference the quantum state must be comply with the local (and global) particle number SSR. This is essentially the same approach as in [64], [53], [4] and in the present paper. For two mode systems the *observable* system density operator $\widetilde{\widehat{\rho}}$ is obtained from the density operator $\widehat{\rho}$ that would apply *if* such a phase reference existed via the expression

$$\widetilde{\widehat{\rho}} = \sum_{n_A n_B} \widehat{\Pi}_{n_A n_B} \widehat{\rho} \, \widehat{\Pi}_{n_A n_B} = \sum_{n_A n_B} \widehat{\rho}^{(n_A n_B)} \tag{136}$$

where $\widehat{\Pi}_{n_A n_B} = \widehat{\Pi}^2_{n_A n_B}$ is a projector onto sub-system states with n_A, n_B particles in modes A, B respectively. Note that $\widehat{\rho}^{(n_A n_B)} = \widehat{\Pi}_{n_A n_B} \widehat{\rho} \, \widehat{\Pi}_{n_A n_B}$ is not normalised to unity. In fact the probability that there are n_A, n_B particles in modes A, B respectively is given by $P_{n_A n_B} = Tr(\widehat{\Pi}_{n_A n_B} \widehat{\rho} \, \widehat{\Pi}_{n_A n_B}) = Tr(\widehat{\rho}^{(n_A n_B)})$, so $Tr(\widetilde{\widehat{\rho}}) = \sum_{n_A n_B} P_{n_A n_B} = 1$. For *separable* states defined here as in Eq.(3), the expression in (136) for the density operator is the *same* as that used here, since with $\widehat{\rho}$ given by Eq.(3) and with $\sum_{n_A n_B} \widehat{\Pi}_{n_A n_B} (\widehat{\rho}^A_R \otimes \widehat{\rho}^B_R) \widehat{\Pi}_{n_A n_B} = \widehat{\rho}^A_R \otimes \widehat{\rho}^B_R$ it is easy to show that $\widetilde{\widehat{\rho}} = \widehat{\rho}_{sep}$. For general mixed states Wiseman et al introduce in Ref. [111] the idea of *particle entanglement* by defining its measure $E_P(\widehat{\rho})$ by

$$E_P(\widehat{\rho}) = \sum_{n_A n_B} P_{n_A n_B} E_M(\widehat{\rho}^{(n_A n_B)}) = E_P(\widetilde{\widehat{\rho}}) \tag{137}$$

where $E_M(\widehat{\rho}^{(n_A n_B)})$ is a measure of the *mode entanglement* associated with the (unnormalised) state $\widehat{\rho}^{(n_A n_B)}$. This might be taken as the *entropy* of mode entanglement $E_M(\widehat{\sigma}) = -Tr(\widehat{\sigma}_A \ln \widehat{\sigma}_A)$ for normalised density operators $\widehat{\sigma}$, where the reduced density operator for mode A is $\widehat{\sigma}_A = Tr_B(\widehat{\sigma})$. Note that from $\widehat{\Pi}_{n_A n_B} \widehat{\Pi}_{m_A m_B} = \delta_{n_A m_A} \delta_{n_B m_B} \widehat{\Pi}_{n_A n_B}$ the particle entanglement measure $E_P(\widehat{\rho})$ is the same for $\widetilde{\widehat{\rho}}$, the observable density operator for the system. In the case of the separable state for modes A, B given in (3) it is straightforward to show that

$$\widehat{\rho}^{(n_A n_B)}_{sep} = \sum_R P_R (\widehat{\Pi}_{n_A} \widehat{\rho}^A_R \widehat{\Pi}_{n_A}) \otimes (\widehat{\Pi}_{n_B} \widehat{\rho}^B_R \widehat{\Pi}_{n_B}) \tag{138}$$

$$P^{sep}_{n_A n_B} = \sum_R P_R P_{n_A}(\widehat{\rho}^A_R) P_{n_B}(\widehat{\rho}^B_R) \tag{139}$$

where $\widehat{\Pi}_{n_A}$ and $\widehat{\Pi}_{n_B}$ are projectors onto sub-system states in modes A, B respectively with n_A and n_B particles in the respective modes ($\widehat{\Pi}_{n_A n_B} = \widehat{\Pi}_{n_A} \otimes \widehat{\Pi}_{n_B}$), with $P_{n_A}(\widehat{\rho}_R^A) = Tr_A(\widehat{\Pi}_{n_A} \widehat{\rho}_R^A)$ and $P_{n_B}(\widehat{\rho}_R^B) = Tr_B(\widehat{\Pi}_{n_B} \widehat{\rho}_R^B)$ being the probabilities of finding n_A and n_B particles in the respective modes when the corresponding sub-system states are $\widehat{\rho}_R^A$ and $\widehat{\rho}_R^B$. Since the state $\widehat{\rho}_{sep}^{(n_A n_B)}$ is clearly a separable state of the form (3) for the modes A, B, the corresponding measure of mode entanglement must be zero. It then follows from the general expression (137) that the particle entanglement measure is also zero for the separable state. This is as expected.

$$E_P(\widehat{\rho}_{sep}) = 0 \tag{140}$$

For the pure states considered in [39] we note that among them is the two boson state $|1\rangle_A \otimes |1\rangle_B$ which has one boson in each of the two modes A, B. The particle entanglement measure $E_P(\widehat{\rho})$ is zero for this state (where $\widehat{\rho} = (|1\rangle \langle 1|)_A \otimes (|1\rangle \langle 1|)_B$), consistent with it being a separable rather than an entangled state. This indicates that Wiseman et al [39] do *not* consider that entanglement occurs due to symmetrization, as the first quantization form for the state might indicate. However, finding $E_P(\widehat{\rho})$ to be zero does not always shows that the state is separable, as the case of the *relative phase state* (defined in Appendix **J** of paper II, see also [32]) shows. As is shown there, $E_P(\widehat{\rho}) = 0$ for the relative phase state, yet the state is clearly an entangled one. Just as some entangled states have zero spin squeezing, some entangled states may be associated with a zero particle entanglement measure. Nevertheless a non-zero result for the particle entanglement measure $E_P(\widehat{\rho})$ *shows* that the state must be *entangled* - again we have a *sufficiency* test. However, as in the case of other entanglement measures the problem with using the particle entanglement measure to detect entangled states is that there is no obvious way to measure it experimentally.

3.4.5   General Form of Non-Entangled States

To summarise: basically the sub-systems are *single modes* that the identical bosons can occupy, the super-selection rule for identical bosons, massive or otherwise, prohibits states which are coherent superpositions of states with different numbers of bosons, and the only physically allowable $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ..for the separate mode sub-systems that are themselves compatiible with the local particle number SSR are allowed. For single mode sub-systems these can be written as statistical mixtures of states with definite numbers of bosons in the form

$$\widehat{\rho}_R^A = \sum_{n_A} P_{n_A}^A |n_A\rangle \langle n_A| \qquad \widehat{\rho}_R^B = \sum_{n_B} P_{n_B}^B |n_B\rangle \langle n_B| \qquad .. \tag{141}$$

However, in cases where the sub-systems are *pairs of modes* the density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ..for the separate sub-systems are still required to conform to the symmetrisation principle and the super-selection rule. The forms for $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, .. are now of course more complex, as entanglement *within* the pairs of modes A_1, A_2 associated with sub-system A, the pairs of modes B_1, B_2

associated with sub-system B, etc is now possible within the definition for the general non-entangled state Eq. (3) for these *pairs* of modes. Within each pair of modes A_1, A_2 statistical mixtures of states with differing total numbers N_A bosons in the two modes are possible and the sub-system density operators are based on states of the form given in Eq. (102). We have

$$
\begin{aligned}
|\Phi_{N_A}\rangle_A &= \sum_{k=0}^{N_A} C_{A\Phi}(N_A, k)\, |k\rangle_{A_1} \otimes |N_A - k\rangle_{A_2} \\
\widehat{\rho}_R^A &= \sum_{N_A=0}^{\infty} \sum_{\Phi} P_{\Phi N_A}\, |\Phi_{N_A}\rangle_A \langle \Phi_{N_A}|_A
\end{aligned}
\tag{142}
$$

with analogous expressions for the density operators $\widehat{\rho}_R^B$ etc for the other pairs of modes. Note that $|\Phi_{N_A}\rangle_A$ only involves quantum superpositions of states with the same total number of bosons N_A. The expression (**208**) in Appendix **B** of paper II is of this form.

3.5   Bipartite Systems

We now consider the bipartite case where there are just two sub-systems involved. The simplest case is where each sub-system involves only a single mode, such as for two modes in a double well potential when only a single hyperfine state is involve. Another important case is where each sub-system contains two modes, such as in the double well case where modes with two different hyperfine states are involved.

3.5.1   Two Single Modes - Coherence Terms

The general non-entangled state for modes $\widehat{a}$ and $\widehat{b}$ is given by

$$
\widehat{\rho} = \sum_R P_R\, \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B
\tag{143}
$$

and as a consequence of the requirement that $\widehat{\rho}_R^A$ and $\widehat{\rho}_R^B$ are allowed quantum states for modes $\widehat{a}$ and $\widehat{b}$ satisying the super-selection rule, it follows that

$$
\begin{aligned}
\langle (\widehat{a})^n \rangle_a &= Tr(\widehat{\rho}_R^A (\widehat{a})^n) = 0 & \langle (\widehat{a}^\dagger)^n \rangle_a &= Tr(\widehat{\rho}_R^A (\widehat{a}^\dagger)^n) = 0 \\
\langle (\widehat{b})^m \rangle_b &= Tr(\widehat{\rho}_R^B (\widehat{b})^m) = 0 & \langle (\widehat{b}^\dagger)^m \rangle_b &= Tr(\widehat{\rho}_R^B (\widehat{b}^\dagger)^m) = 0
\end{aligned}
\tag{144}
$$

Thus coherence terms are zero. As we will see these results will limit spin squeezing to entangled states of modes $\widehat{a}$ and $\widehat{b}$. Note that similar results also apply when non-entangled states for the original modes $\widehat{c}$ and $\widehat{d}$ are considered - $\langle (\widehat{c})^n \rangle_c = 0$, etc..

3.5.2 Two Pairs of Modes - Coherence Terms

In this case the general non-entangled state where A and B are pairs of modes - $\widehat{a}_1$, $\widehat{a}_2$ associated with sub-system A, and modes $\widehat{b}_1$, $\widehat{b}_2$ associated with sub-system B, the overall density operator is of the form (143). Consistent with the requirement that the sub-system density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$ conform to the symmetrisation principle and the super-selection rule, these density operators will *not* in general represent separable states for their single mode sub-systems $\widehat{a}_1$, $\widehat{a}_2$ or $\widehat{b}_1$, $\widehat{b}_2$ - and may even be entangled states. As a result when considering *non-entangled* states for the sub-systems A and B we now have

$$\left\langle (\widehat{a}_i^\dagger \widehat{a}_j)^n \right\rangle_A = Tr(\widehat{\rho}_R^A (\widehat{a}_i^\dagger \widehat{a}_j)^n) \neq 0 \qquad i,j = 1,2$$

$$\left\langle (\widehat{b}_i^\dagger \widehat{b}_j)^n \right\rangle_A = Tr(\widehat{\rho}_R^B (\widehat{b}_i^\dagger \widehat{b}_j)^n) \neq 0 \qquad i,j = 1,2 \tag{145}$$

in general. In this case where the sub-systems are *pairs* of modes the spin squeezing entanglement tests as in **Eqs.(50) - (52) in paper II** for sub-systems consisting of *single* modes cannot be applied, as we will see. Nevertheless, there are still tests of bipartite entanglement involving spin operators.

4 Discussion and Summary of Key Results

This paper is mainly concerned with two mode entanglement for systems of identical massive bosons, though multimode entanglement is also considered. These bosons may be atoms or molecules as in cold quantum gases. In the present paper we focus on the definition and general features of entanglement, whilst in the accompanying paper we consider spin squeezing and other tests for entanglement.

The present paper starts with the *general definition* of *entanglement* for a system consisting of several *sub-systems*, and highlights the distinctive *features* of entangled states in regard to measurement probabilities for joint measurements on the sub-systems in contrast to the results for non-entangled or *separable* states. The relationship between entanglement and *hidden variable theory* is then explored followed by a discussion of key *paradoxes* such as EPR and violations of Bell inequalities. The notion of entanglement *measures* and entanglement *tests* was briefly introduced, the latter being covered more fully in the accompanying Paper II.

The paper then focuses on entanglement for systems of identical massive particles in the regime of non-relativistic quantum physics. A careful analysis is first given regarding the proper definition of a non-entangled state for systems of identical particles, and hence by implication the proper definition of an entangled state. Noting that entanglement is meaningless until the subsystems being entangled are specified, it is pointed out that whereas it is not possible to distinguish identical particles and hence the individual particles are not legitimate sub-systems, the same is not the case for the single particle states or modes, so the *modes* are then the the rightful *sub-systems* to be considered as being entangled or not. In this approach where the sub-systems are modes, situations where there are differing numbers of identical particles are treated as different quantum states, not as differing physical systems, and the *symmetrisation principle* required of quantum states for identical particle systems will be satisfied by using Fock states to describe the states.

Furthermore, it is argued that the overall quantum states should conform to the superselection rule that excludes quantum superposition states of the form (99) as allowed quantum states for systems of identical particles - massive or otherwise. Although the justication of the SSR in terms of observers and their *reference frames* formulated by other authors has also been presented for completeness, a number of fairly *straightforward reasons* were given for why it is appropriate to apply this superselection rule for massive bosons, which may be summarised as: 1. No way is known for creating such states; 2 No way is known for measuring all the properties of such states, even if they existed; and 3. There is no need to invoke the existence of such states in order to understand coherence and interference effects. Invoking the existence of states that as far as we know cannot be made or measured, and for which there are no known physical effects that require their presence seems a rather unnecessary feature to add to the non-relativistic quantum physics of many body systems, and considerations based on the general principle of simplicity (Occam's razor)

would suggest not doing this until a clear physical justification for including them is found. As two mode fermionic systems are restricted to states with at most two fermions, the focus of the paper is then on bosonic systems, where large numbers of bosons can occupy two mode systems.

However, although there is related work involving local particle number super-selection rules, *this paper differs* from a number of others by *extending* the *super-selection rule* to also apply to the density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... for the *mode sub-systems A, B, ...* that occur in the definition (3) of a *general non-entangled* state for systems of identical particles. Hence it follows that the definition of *entangled states* will differ in this paper from that which would apply if density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... allowed for coherent superpositions of number states within each mode. In fact more states are regarded as entangled in terms of the definition in the present paper. Indeed, if *further restrictions* are placed on the sub-system density operators - such as requiring them to specify a fixed number of bosons - the set of entangled states is further enlarged. The *simple justification* for our viewpoint on applying the *local particle number super-selection rule* has three aspects. Firstly, since experimental arrangements in which only one bosonic mode is involved can be created, the same reasons (see last paragraph) justify applying the super-selection rule to this mode system as applied for the system as a whole. Secondly, measurements can be carried out on the separate modes, and the joint probability for the outcomes of these measurements determined. For a non-entangled state the joint probability (23) for these measurements depends on all the density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... for the mode sub-systems as well as the probability P_R for the product state $\widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes$ ... occuring when the general mixed non-entangled state is prepared, which can be accomplished by local preparations and classical communication. For the non-entangled state the form of the joint probability $P_{AB..}(i, j, ...)$ for measurements on all the sub-systems is given by the products of the individual sub-system probabilities $P_A^R(i) = Tr(\widehat{\Pi}_i^A \widehat{\rho}_R^A)$, etc that measurements on the sub-systems $A, B, ...$ yield the outcomes λ_i^A etc when the sub-systems are in states $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... , the overall products being weighted by the probability P_R that a particular product state is prepared. If $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... did not represent allowed quantum states then the interpretation of the joint probability as this statistical average would be unphysical Thirdly, attempts to allow the density operators $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, ... for the mode sub-systems to violate the super-selection rule provided that the reduced density operators $\widehat{\rho}_A$, $\widehat{\rho}_B$ for the separate modes are consistent with it are shown not to be possible in general.

As well as the above justifications for applying the super-selection rule to both the overall multi-mode state for systems of identical particles and the separate sub-system states in the definition of non-entangled states, a more sophisticated justification based on considering SSR to be the consequence of describing the quantum state by an observer (Charlie) whose phase reference is unknown has also been presented in detail in Appendix 10 for completeness. For the sub-systems *local reference frames* are involved. The SSR is seen as a special case of a general SSR which forbids quantum states from exhibiting coherences

between states associated with *irreducible representations* of the transformation group that relates reference frames, and which may be the *symmetry group* for the system.

In regard to entanglement measures, we discussed the particle entanglement measure of Wiseman et al [39], [40] and found that a non-zero result for the particle entanglement measure *shows* that the state must be *entangled.* However, as for other entanglement measures the problem with using the particle entanglement measure to detect entangled states is that there is no obvious way to measure it experimentally. On the other hand, as will be seen in the accompanying Paper II, the quantities involved in entanglement tests can be measured experimentally.

References

[1] Einstein, A., Podolsky, B. and Rosen, N., *Phys. Rev.* **1935**, *47,* 777.

[2] Schrodinger, E., *Naturwiss.* **1935**, *23*, 807. Translated in *Quantum Theory and Measurement;* eds. J. A. Wheeler and W. Zurek, Princeton, N. J.: 1983.

[3] Bell, J. S., *Physics* **1964**, *1*, 195.

[4] Dalton, B. J., Heaney, L., Goold, J. Garraway, B. M. and Busch, Th., *New J. Phys.* **2014**, *16*, 013026.

[5] Werner, R. F., *Phys. Rev. A* **1989**, *40,* 4277.

[6] Brunner, N., Cavalcanti, D., Pironio, S. Scarani, V. and Wehner, S., *Rev. Mod. Phys.* **2014**, *86*, 419.

[7] Hensen, B., Bernien, H., *et al, Nature* **2015**, *526*, 682.

[8] von Neumann, J., *Mathematical Foundations of Quantum Mechanics:* Springer-Verlag: Berlin, **1932**.

[9] Wheeler, J. A. and Zurek, W. H., *Quantum Theory and Measurement;* Princeton University Press: Princeton, N.J., **1983**.

[10] Zurek, W. H., *Phys. Rev. D* **1981**, *24*, 1516.

[11] Peres, A. , *Quantum Theory: Concepts and Methods;* Kluwer: Dortrecht, **1993**.

[12] Nielsen, M. A. and Chuang, I. L., *Quantum Computation and Quantum Information;* Cambridge: Cambridge, **2000**.

[13] Vedral, V. , *Introduction to Quantum Information Science;* Oxford: Oxford, **2007**.

[14] Barnett, S. M., *Quantum Information;* Oxford: Oxford, **2009**.

[15] Reid, M. D., Drummond, P. D., Bowen, W. P., Cavalcanti. E. G. , Lam., P. K., Bachor, H. A., Anderson, U. L. and Leuchs, G., *Rev. Mod. Phys.* **2009**, *81*, 1727.

[16] Reid, M. D., He, Q.-Y. and Drummond, P. D., *Front. Phys.* **2012**, *7*, 72.

[17] Gisin N., *Phys. Letts. A,* **1991,** *154,* 201.

[18] Barrett, J., *Phys. Rev. A* **2002,** *65,* 043302.

[19] Toth, G. and Acin, A., *Phys. Rev. A* **2006,** *74,* 030306(R).

[20] Almeida, M. L., Pironio, S., Barrett, J., Toth, G. and Acin, A., *Phys. Rev. Letts.* **2007,** *99,* 040403.

[21] Ollivier, H. and Zurek, W. H., *Phys. Rev. Letts.* **2002,** *88,* 017901.

[22] Modi, K, Brodutch, A., Cable, H., Paterek, T. and Vedral, V., *Rev. Mod. Phys.* **2012**, *84*, 1655.

[23] Wiseman, H. M., Jones, S. J. and Doherty, A. C., *Phys. Rev. Letts.* **2007,** *98,* 140402.

[24] Jones, S. J., Wiseman, H. M. and Doherty, A. C. , *Phys. Rev. A* **2007,** *76,* 052116.

[25] Cavalcanti, E., Jones, S. J., Wiseman, H. M. and Reid, M. D., *Phys. Rev. A* **2009,** *80,* 032112.

[26] He, Q. Y. Gong, Q. H. and Reid, M. D., *Phys. Rev. Letts.* **2015,** *114,* 060402.

[27] Simon, C. *Phys. Rev. A* **2002,** *66,* 052323.

[28] Hines, A.P., McKenzie, R.H. and Milburn, G.J., *Phys. Rev. A* **2003**, *67,* 013609.

[29] Terra Cunha, M.O., Dunningham, J.A. and Vedral, V., *Proc. Roy. Soc. A* **2007,** *463,* 2277.

[30] Horodecki, R., Horodecki, P., Horodecki, M. and Horodecki, K., *Rev. Mod. Phys.* **2009**, *81*, 865.

[31] Guhne, O. and Toth, G., *Phys. Rep.* **2009**, *474,* 1.

[32] Dalton, B.J. and Ghanbari, S. *J. Mod. Opt.* **2012**, *59,* 287, *ibid* **2013**, *60,* 602.

[33] Goold, J., Heaney, L., Busch, Th. and Vedral V., *Phys. Rev. A* **2009**, *80,* 022338.

[34] Tichy, M. C., Mintert, F. and Buchleitner, A. *J. Phys. B: At. Mol. Opt. Phys.* **2011**, *44,* 192001.

[35] Amico, L., Fazio, R., Osterloh, A. and Vedral, V., *Rev. Mod. Phys.* **2008,** *80***,** 517.

[36] Islam, R., Ma, R., Preiss, P. M., Tai, M. E., Lukin, A., Rispoli, M. and Greiner M., *Nature* **2015**, *528,* 77.

[37] Cramer, M., Bernard, A., Fabbri, N., Fallani, L., Fort, C., Rosi, S., Caruso, F., Inguscio M. and Plenio, M. B., *Nat. Comm.* **2013**, *4:2161.*

[38] Daley, A. J., Pichler, H., Schachenmayer, J. and Zoller, P. *Phys. Rev. Letts.* **2012,** *109,* 020505.

[39] Wiseman, H. M. and Vaccaro, J. A., *Phys. Rev. Letts.* **2003,** *91,* 097902.

[40] Dowling, M. R., Doherty, A. C. and Wiseman, H. M., *Phys. Rev. A* **2006**, *73, 052323.*

[41] Peres, A., *Phys. Rev. Letts.* **1996,** *77,* 1413.

[42] Horodecki, M., Horodecki, P., and Horodecki, R., *Phys. Letts. A.* **1996**, *223,* 1.

[43] Horodecki, P., *Phys. Letts. A.* **1997**, *232,* 333.

[44] Benatti, F., Floreanni, R. and Marzolino, U., **2010**, *Ann. Phys.,325,* 924.

[45] Benatti, F., Floreanni, R. and Marzolino, U., **2011**, *J. Phys. B: At. Mol. Opt. Phys.,44,* 1001.

[46] Lunkes, C., Brukner, C. and Vedral, V., *Phys. Rev. Letts.* **2005,** *95,* 030503.

[47] Bartlett, S. D., Doherty, A. C., Spekkens, R. W. and Wiseman, H. W., *Phys. Rev. A* **2006**, *73,* 022311.

[48] Masanes, L., Liang, Y.-C. and Doherty, A. C., *Phys. Rev. Letts.* **2008,** *100,* 090403.

[49] Verstraete, F. and Cirac, J. I. ,*Phys. Rev. Letts.* **2003,** *91,* 010404.

[50] Schuch, N., Verstraete, F. and Cirac, J. I., *Phys. Rev. A* **2004**, *70,* 042310.

[51] Dirac, P. A. M., *Quantum Mechanics,* Oxford: Oxford, **1930**.

[52] Wick, G. C., Wightman, A. S. and Wigner, E. P., *Phys. Rev.* **1952**, *88,* 101.

[53] Bartlett, S. D., Rudolph, T. and Spekkens, R. W. ,*Rev. Mod. Phys.* **2007,** *79***,** 555.

[54] Hillery, M. and Zubairy, M.S., *Phys. Rev. Letts.* **2006,** *96,* 050503.

[55] Hillery, M., Dung, H. T. and Niset, J. *Phys. Rev. A* **2009,** *80,* 052335.

[56] Vaccaro, J. A. , Anselmi, F., Wiseman, H. M. and Jacobs, K., *Phys. Rev. A* **2008**, *77,* 032114.

[57] White, G. A., Vaccaro, J. A. and Wiseman, H. M., *Phys. Rev. A* **2009**, *79,* 032109.

[58] Paterek, T., Kurzynski, P., Oi, D. K. L. and Kaszlikowski, D., *New J. Phys.* **2011**, *13,* 043027.

[59] Aharonov, Y. and Susskind, L., *Phys. Rev.* **1967,** *155,* 1428.

[60] Bartlett, S. D. and Wiseman, H. W., *Phys. Rev. Letts.* **2003,** *91,* 097903.

[61] Sanders, B. C., Bartlett, S. D., Rudolph, T. and Knight, P. L., *Phys. Rev. A* **2003,** *68,* 042329.

[62] Kitaev, A., Mayers, D. and Preskill, J. , *Phys. Rev. A* **2004**, *69,* 052326.

[63] van Enk, S. J., *Phys. Rev. A* **2005**, *72,* 064306.

[64] Bartlett, S. D., Rudolph, T. and Spekkens, R. W. ,*Int. J. Quant. Infn.* **2006**, *4,* 17.

[65] Wodkiewicz, K. and Eberly, J. H., *J. Opt. Soc. Amer. B* **1985**, *2,* 458.

[66] Kitagawa, M. and Ueda, M., *Phys. Rev. A* **1993**, *47,* 5138.

[67] Sorensen, A., Duan, L.-M., Cirac, J.I. and Zoller, P. *Nature* **2001,** *409,* 63.

[68] Micheli, A., Jaksch, D., Cirac, J. I. and Zoller, P., *Phys. Rev. A* **2003**, *67,* 013607.

[69] Toth, G., Knapp, C., Guhne, O. and Briegel, H. J., *Phys. Rev. Letts.* **2007,** *99,* 250405.

[70] Hyllus, P., Pezze, L., Smerzi, A. and Toth, G. *Phys. Rev. A* **2012**, *86,* 012237.

[71] Leggett, A.J., *Rev. Mod. Phys.* **2001,** *73,* 307.

[72] Pegg, D. T., Barnett, S. M. and Jeffers, J., *J. Mod. Opt.* **2002**, *49,* 913.

[73] Caves, C. M., Fuchs, C. A. and Schack, R. *Phys. Rev. A* **2002**, *65,* 022305.

[74] Pusey, M. F., Barrett, J. and Rudolph, T. *Nature Phys.* **2012**, *8,* 475.

[75] Pegg, D. T. and Jeffers, J., *J. Mod. Opt.* **2005**, *52,* 1835.

[76] Reid, M. D., *Phys. Rev. A* **1989**, *40,* 913.

[77] Cavalcanti, E.G., Drummond, P.D., Bachor, H.A. and Reid, M. D., *Opt. Exp.* **2009**, *17,* 18694.

[78] Spreeuw, R. J. C., *Found. Phys.* **1998,** *28,* 361.

[79] Borges, C. V. S., Hor-Meyl, M., Huguenin, J. A. O. and Khoury, A. Z. , *Phys. Rev. A* **2010**, *82,* 033833.

[80] Aiello, A., Toppel, F., Marquart, C., Giacobino, E. and Leuchs, G. **2014**, *Classical Entanglement: Oxymoron or Resource ? ArXiv 1409.0213.*

[81] Isham, C. J., *Quantum Theory;* Imperial College Press: London, **1995**.

[82] Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A., *Phys. Rev. Letts.* **1969,** *23,* 880.

[83] Clauser, J. F. and Shimony, A. *Rep. Prog. Phys.,* **1978,** *41,* 1881.

[84] Reid, M. D., *Phys. Rev. A* **2000**, *62,* 022110.

[85] Reid, M. D., **2003**, In *Quantum Squeezing*, P. D. Drummond and Z. Ficek, Eds. (Springer).

[86] Greenberger, D. M., Horne, M. and Zeilinger, A., **2003**, In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, M. Kafatos, Ed. (Kluwer).

[87] Greenberger, D. M., Horne, M., Shimony, A. and Zeilinger, A., *Amer. J. Phys.* **1990,** *58,* 1131.

[88] Mermin, N. D., *Physics Today* **1990**, *43*, 9.

[89] Bohm, D., *Quantum Theory;* Prentice-Hill, New York, **1951**.

[90] Rinner, S. and Werner, E., *Cent. Eur. J. Phys.* **2008**, *6*, 178.

[91] Reid, M. D. and Cavalcanti, E. G. *J. Mod. Opt.* **2005**, *52*, 2245.

[92] Rosales-Zarate, L., Teh, R. Y., Kiesewetter, S. Brolis, A., Ng, K. and Reid, M. D. ,*J. Opt. Soc. Am. B* **2015**, *32,* A82.

[93] Brune, M., Hagley, E., Dreyer, J., Maitre, X., Maali, C., Wunderlich, C., Raimond, J. M. and Haroche, S. *Phys. Rev. Letts.* **1996,** *77,* 4887.

[94] Pezze, L. and Smerzi, A., *Phys. Rev. Letts.* **2009,** *102,* 100401.

[95] Killoran N, Cramer M and Plenio M B., *Phys. Rev. Letts.* **2014,** *112,* 150501.

[96] He, Q. Y., Drummond, P. D., Olsen, M. K. and Reid, M. D., *Phys. Rev. A* **2012,** *86,* 023626.

[97] Heaney, L., Lee, S.-W. and Jaksch, D., *Phys. Rev. A* **2010,** *82,* 042116.

[98] Cable, H., Knight, P. L. and Rudolph, T., *Phys. Rev. A* **2005**, *71*, 042107.

[99] Bach, R. and Rzazewskii, K, *Phys. Rev. A* **2004**, *70*, 063622.

[100] Javanainen, J. and Yoo, S. M., *Phys. Rev. Letts.* **1996,** *76,* 161.

[101] Stenholm, S., *Phys. Scripta* **2002**, *T102*, 89.

[102] Molmer, K., *Phys. Rev. A* **1997**, *55*, 4247.

[103] Mandel, L. and Wolf, E., *Optical Coherence and Quantum Optics,* Cambridge, London, **1995.**

[104] Wiseman, H. W. and Vaccaro, J., *Phys. Rev. A* **2002**, *65*, 043605.

[105] Wiseman, H. W. and Vaccaro, J., *Phys. Rev. A* **2002**, *65*, 043606.

[106] Rempe, G., Walther, H. and Klein, N., *Phys. Rev. Letts.* **1996,** *58*, 353.

[107] Brune, M., Schmidt-Kaler, F., Maaili, A., Dreyer, J., Hagley, E., Raimond, J. M. and Haroche, S., *Phys. Rev. Letts.* **1996,** *76,* 1800.

[108] McNeil, K. J. and Gardiner, C. W., *Phys. Rev. A* **1983**, *28*, 1560.

[109] Ou, Z. Y., Pereira, S. F., Kimble, H. J. and Peng, K. C., *Phys. Rev. Letts.* **1992,** *68,* 3663.

[110] Wu, L.-A., Kimble, H. J., Hall, J. L. and Wu, H., *Phys. Rev. Letts.* **1986,** *57,* 2520.

[111] Dowling, M. R., Bartlett, S. D., Rudolph, T. and Spekkens, R. W., *Phys. Rev. A* **2006**, *74*, 052113.

[112] Dunningham, J. A., Rico Gutierrez, L. M. and Palge, V., *Optics and Spectroscopy,* **2011**, *111*, 528.

[113] Tinkham, M., *Group Theory and Quantum Mechanics;* McGraw-Hill, New York, **1964**.

5 Acknowledgements

6 Appendix 1 - Projective Measurements and Conditional Probabilities

6.0.3 Projective Measurements

For simplicity, we will only consider *projective* (or von Neumann) measurements rather than more general measurements involving *positive operator measurements* (POM). If $\widehat{\Omega}$ is a physical quantity associated with the system, with eigenvalues λ_i and with $\widehat{\Pi}_i$ the projector onto the subspace with eigenvalue λ_i then the probability $P(i)$ that measurement of $\widehat{\Omega}$ leads to the value λ_i is given by [81]

$$P(i) = Tr(\widehat{\Pi}_i \widehat{\rho}) \tag{146}$$

For projective measurements $\widehat{\Pi}_i = \widehat{\Pi}_i^2 = \widehat{\Pi}_i^\dagger$ and $\sum_i \widehat{\Pi}_i = 1$, together with $\widehat{\Omega}\widehat{\Pi}_i = \widehat{\Pi}_i\widehat{\Omega} = \lambda_i\widehat{\Pi}_i$.

Following the measurement which leads to the value λ_i the density operator is different and given by

$$\widehat{\rho}_{cond}(\widehat{\Omega}, i) = (\widehat{\Pi}_i \widehat{\rho} \widehat{\Pi}_i)/P(i) \tag{147}$$

This is known as the *reduction of the wave function*, and can be viewed in two ways. From an ontological point of view a quantum projective measurement *changes* the quantum state significantly because the interaction with the measurement system is not just a small perturbation, as it can be in classical physics. From the epistomological point of view we know what value the physical quantity $\widehat{\Omega}$ now has, so if measurement of $\widehat{\Omega}$ were to be *repeated* immediately it would be expected – with a probability of unity - that the value would be λ_i. The new density operator $\widehat{\rho}_{cond}(\widehat{\Omega}, i)$ satisfies this requirement. It also satisfies the standard requirements of Hermitiancy, unit trace, positivity - as is easily shown.

To show this formally we have for the *mean value* for $\widehat{\Omega}$ following the measurement

$$\begin{aligned}
\left\langle \widehat{\Omega} \right\rangle_i &= Tr(\widehat{\Omega}\,\widehat{\rho}_{cond}(\widehat{\Omega}, i)) \\
&= Tr(\widehat{\Omega}\,(\widehat{\Pi}_i\widehat{\rho}\widehat{\Pi}_i))/P(i) \\
&= \lambda_i Tr(\widehat{\Pi}_i\widehat{\rho})/P(i) \\
&= \lambda_i
\end{aligned} \tag{148}$$

whilst for the *variance*

$$\begin{aligned}
\left\langle \Delta\widehat{\Omega}^2 \right\rangle_i &= Tr((\widehat{\Omega} - \left\langle \widehat{\Omega} \right\rangle_i)^2\,\widehat{\rho}_{cond}(\widehat{\Omega}, i)) \\
&= Tr(\widehat{\Omega}^2\,\widehat{\rho}_{red}(i)) - \left\langle \widehat{\Omega} \right\rangle_i^2 \\
&= \lambda_i^2 - \lambda_i^2 \\
&= 0
\end{aligned} \tag{149}$$

which is zero as expected.

If following the measurement of $\widehat{\Omega}$ the results of the measurement were discarded then the density operator after the measurement is

$$\widehat{\rho}_{cond}(\widehat{\Omega}) = \sum_i P(i)\,\widehat{\rho}_{cond}(\widehat{\Omega}, i) = \sum_i \widehat{\Pi}_i \widehat{\rho} \widehat{\Pi}_i \tag{150}$$

which is the sum of the $\widehat{\rho}_{cond}(\widehat{\Omega}, i)$ each weighted by the probability $P(i)$ of the result λ_i occuring. Note that the expression for $\widehat{\rho}_{cond}(\widehat{\Omega})$ is not the same as the original density operator $\widehat{\rho}$. This is to be expected from both the epistimological and ontological points of view, since although we do *not* know *what* value λ_i has occurred, it is known that *a* definite value for $\widehat{\Omega}$ has been found, or that measurement process has destroyed any *coherences* that previously existed between different eigenstates of $\widehat{\Omega}$. We note that $\widehat{\rho}_{cond}(\widehat{\Omega})$ also satisfies the standard requirements of Hermitiancy, unit trace, positivity - as is easily shown.

6.0.4 Conditional Probabilities

Suppose we follow the measurement of $\widehat{\Omega}$ resulting in eigenvalue λ_i with a measurement of $\widehat{\Lambda}$ resulting in eigenvalue μ_j where the projector associated with the latter measurement is $\widehat{\Xi}_j$. Then the *conditional probabiltity* of measuring $\widehat{\Lambda}$ resulting in eigenvalue μ_j following the measurement of $\widehat{\Omega}$ that resulted in eigenvalue λ_i would be

$$\begin{aligned} P(j|i) &= Tr(\widehat{\Xi}_j \widehat{\rho}_{cond}(\widehat{\Omega}, i)) \\ &= Tr(\widehat{\Xi}_j (\widehat{\Pi}_i \widehat{\rho} \widehat{\Pi}_i))/P(i) \\ &= Tr((\widehat{\Xi}_j \widehat{\Pi}_i)\,\widehat{\rho}\,(\widehat{\Pi}_i \widehat{\Xi}_j))/P(i) \end{aligned} \tag{151}$$

where the cyclic properties of the trace and the idempotent property of the projector have been used. If the measurements had taken place in the reverse order the conditional probabiltity of measuring $\widehat{\Omega}$ resulting in eigenvalue λ_i following the measurement of $\widehat{\Lambda}$ that resulted in eigenvalue μ_j would be

$$P(i|j) = Tr((\widehat{\Pi}_i \widehat{\Xi}_j)\,\widehat{\rho}\,(\widehat{\Xi}_j \widehat{\Pi}_i))/P(j) \tag{152}$$

We note that the actual probability of measuring λ_i then μ_j would be the *joint probability*

$$P(j\ after\ i) = P(j|i)\,P(i) = Tr((\widehat{\Xi}_j \widehat{\Pi}_i)\,\widehat{\rho}\,(\widehat{\Pi}_i \widehat{\Xi}_j)) \tag{153}$$

whilst the actual probability of measuring μ_j then λ_i would be the joint probability

$$P(i\ after\ j) = P(i|j)\,P(j) = Tr((\widehat{\Pi}_i \widehat{\Xi}_j)\,\widehat{\rho}\,(\widehat{\Xi}_j \widehat{\Pi}_i)) \tag{154}$$

and we note that in general these two joint probabilities are different.

If however, the two physical quantities *commute*, then there are a complete set of simultaneous eigenvectors $\left|\lambda_i, \mu_j\right\rangle$ for $\widehat{\Omega}$ and $\widehat{\Lambda}$. It is then straightforward

to show that $\widehat{\Pi}_i\widehat{\Xi}_j = \widehat{\Xi}_j\widehat{\Pi}_i$, in which case $P(j\ after\ i) = P(i\ after\ j) = P(i,j)$, so it does not matter which order the measurements are carried out. The overall result

$$
\begin{aligned}
P(i,j) &= P(j|i)\,P(i) = P(i|j)\,P(j) \\
&= Tr(\widehat{\Pi}_i\widehat{\Xi}_j\,\widehat{\rho}\,\widehat{\Xi}_j\widehat{\Pi}_i) \\
&= Tr(\widehat{\Pi}_i\widehat{\Xi}_j\,\widehat{\rho})
\end{aligned}
\tag{155}
$$

is an expression of *Bayes theorem.*

A case of particular importance where this occurs is in situations involving two or more distinct sub-systems, in which the operators $\widehat{\Omega}$ and $\widehat{\Lambda}$ are associated with different sub-systems. For two sub-systems A and B the operators $\widehat{\Omega}$ and $\widehat{\Lambda}$ are of the form $\widehat{\Omega}_A$ and $\widehat{\Omega}_B$, or more strictly $\widehat{\Omega}_A \otimes \widehat{1}_B$ and $\widehat{1}_A \otimes \widehat{\Omega}_B$. It is easy to see that $(\widehat{\Omega}_A \otimes \widehat{1}_B)(\widehat{1}_A \otimes \widehat{\Omega}_B) = \widehat{\Omega}_A \otimes \widehat{\Omega}_B = (\widehat{1}_A \otimes \widehat{\Omega}_B)(\widehat{\Omega}_A \otimes \widehat{1}_B)$, so the operators commute and results such as in Bayes theorem (155) apply.

6.0.5 Conditional Mean and Variance

To determine the *conditioned mean value* of $\widehat{\Lambda}$ after measurement of $\widehat{\Omega}$ has led to the eigenvalue λ_i we use $\widehat{\rho}_{cond}(\widehat{\Omega}, i)$ rather than $\widehat{\rho}$ in the mean formula $\left\langle\widehat{\Lambda}\right\rangle = Tr(\widehat{\Lambda}\widehat{\rho})$. Hence

$$
\begin{aligned}
\left\langle\widehat{\Lambda}\right\rangle_i &= Tr(\widehat{\Lambda}\widehat{\rho}_{cond}(\widehat{\Omega}, i)) \\
&= Tr(\widehat{\Lambda}\,(\widehat{\Pi}_i\widehat{\rho}\widehat{\Pi}_i))/P(i)
\end{aligned}
\tag{156}
$$

Now

$$
\widehat{\Lambda} = \sum_j \mu_j\widehat{\Xi}_j
\tag{157}
$$

so that

$$
\begin{aligned}
\left\langle\widehat{\Lambda}\right\rangle_i &= \sum_j \mu_j\, Tr(\widehat{\Xi}_j\,\widehat{\Pi}_i\widehat{\rho}\widehat{\Pi}_i)/P(i) \\
&= \sum_j \mu_j\, Tr(\widehat{\Xi}_j\widehat{\Pi}_i\widehat{\rho}\widehat{\Pi}_i\widehat{\Xi}_j)/P(i) \\
&= \sum_j \mu_j\, P(j|i)
\end{aligned}
\tag{158}
$$

using $\widehat{\Xi}_j = \widehat{\Xi}_j^2$, the cyclic trace properties and Eq.(151). Hence the conditional mean value is as expected, with the conditional probability $P(j|i)$ replacing $P(j)$ in the averaging process.

For the *conditioned variance* of $\widehat{\Lambda}$ after measurement of $\widehat{\Omega}$ has led to the eigenvalue λ_i we use $\widehat{\rho}_{cond}(\widehat{\Omega}, i)$ rather than $\widehat{\rho}$ and the conditioned mean $\left\langle\widehat{\Lambda}\right\rangle_i$

rather than $\left\langle \widehat{\Lambda} \right\rangle$ in the variance formula $\left\langle \Delta\widehat{\Lambda}^2 \right\rangle = Tr((\widehat{\Lambda} - \left\langle \widehat{\Lambda} \right\rangle)^2 \widehat{\rho})$. Hence

$$
\begin{aligned}
\left\langle \Delta\widehat{\Lambda}^2 \right\rangle_i &= Tr((\widehat{\Lambda} - \left\langle \widehat{\Lambda} \right\rangle_i)^2 \widehat{\rho}_{cond}(\widehat{\Omega}, i)) \\
&= Tr((\widehat{\Lambda} - \left\langle \widehat{\Lambda} \right\rangle_i)^2 (\widehat{\Pi}_i \widehat{\rho} \widehat{\Pi}_i))/P(i)
\end{aligned}
\tag{159}
$$

Now

$$
(\widehat{\Lambda} - \left\langle \widehat{\Lambda} \right\rangle_i)^2 = \sum_j (\mu_j - \left\langle \widehat{\Lambda} \right\rangle_i)^2 \widehat{\Xi}_j
\tag{160}
$$

so that

$$
\begin{aligned}
\left\langle \Delta\widehat{\Lambda}^2 \right\rangle_i &= \sum_j (\mu_j - \left\langle \widehat{\Lambda} \right\rangle_i)^2 \, Tr(\widehat{\Xi}_j \, \widehat{\Pi}_i \widehat{\rho} \widehat{\Pi}_i)/P(i) \\
&= \sum_j (\mu_j - \left\langle \widehat{\Lambda} \right\rangle_i)^2 \, P(j|i)
\end{aligned}
\tag{161}
$$

using the same steps as for the conditioned mean. Hence the conditional variance is as expected, with the conditional probability $P(j|i)$ replacing $P(j)$ in the averaging process.

6.1 Detailed Inequalities for EPR Situation

For *separable states* the *conditional probability* that measurement of $\widehat{p}_A$ on sub-system A leads to eigenvalue p_A given that measurement of $\widehat{p}_B$ on sub-system B leads to eigenvalue p_B is obtained from Eq.(30) as

$$
P(\widehat{p}_A, p_A | \widehat{p}_B, p_B) = \sum_R P_R \, P_A^R(\widehat{p}_A, p_A) P_B^R(\widehat{p}_B, p_B) / \sum_R P_R \, P_B^R(\widehat{p}_B, p_B) \tag{162}
$$

where

$$
P_A^R(\widehat{p}_A, p_A) = Tr_A(\widehat{\Pi}_{p_A}^A \, \widehat{\rho}_R^A) \qquad P_B^R(\widehat{p}_B, p_B) = Tr_B(\widehat{\Pi}_{p_B}^B \, \widehat{\rho}_R^B) \tag{163}
$$

are the probabilities for position measurements in the separate sub-systems. The probability that measurement of $\widehat{p}_B$ on sub-system B leads to eigenvalue p_B is

$$
P(\widehat{p}_B, p_B) = \sum_R P_R \, P_B^R(\widehat{p}_B, p_B) \tag{164}
$$

The *mean* result for measurement of $\widehat{p}_A$ for this *conditional* measurement is from Eq.(19)

$$
\begin{aligned}
\langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B} &= \sum_{p_A} p_A \, P(\widehat{p}_A, p_A | \widehat{p}_B, p_B) \\
&= \sum_R P_R \, \langle \widehat{p}_A \rangle_R \, P_B^R(\widehat{p}_B, p_B) / P(\widehat{p}_B, p_B)
\end{aligned}
\tag{165}
$$

where

$$\langle \widehat{p}_A \rangle_R = \sum_{p_A} p_A P_A^R(\widehat{p}_A, p_A) \tag{166}$$

is the *mean* result for measurement of $\widehat{p}_A$ when the sub-system is in state $\widehat{\rho}_R^A$.

The *conditional variance* for measurement of $\widehat{p}_A$ for the conditional measurement of $\widehat{p}_B$ on sub-system B which led to eigenvalue p_B is from Eq.(20)

$$
\begin{aligned}
\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B, p_B} &= \sum_{p_A} (p_A - \langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B})^2 \, P(\widehat{p}_A, p_A | \widehat{p}_B, p_B) \\
&= \sum_R P_R \, \langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B, p_B}^R \, P_B^R(\widehat{p}_B, p_B) / P(\widehat{p}_B, p_B) \tag{167}
\end{aligned}
$$

where

$$\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B, p_B}^R = \sum_{p_A} (p_A - \langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B})^2 \, P_A^R(\widehat{p}_A, p_A)$$

is a variance for measurement of $\widehat{p}_A$ for when the sub-system is in state $\widehat{\rho}_R^A$ but now with the fluctuation about the mean $\langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B}$ for measurements conditional on measuring $\widehat{p}_B$.

However, for each sub-system state R the quantity $\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B, p_B}^R$ is *minimised* if $\langle \widehat{p}_A \rangle_{\widehat{p}_B, p_B}$ is replaced by the unconditioned mean $\langle \widehat{p}_A \rangle_R$ just determined from $\widehat{\rho}_R^A$. Thus we have an inequality

$$\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B, p_B}^R \geq \langle \Delta \widehat{p}_A^2 \rangle^R \tag{168}$$

where

$$\langle \Delta \widehat{p}_A^2 \rangle^R = \sum_{p_A} (p_A - \langle \widehat{p}_A \rangle)^2 \, P_A^R(\widehat{p}_A, p_A) \tag{169}$$

is the *normal variance* for measurement of $\widehat{p}_A$ for when the sub-system is in state $\widehat{\rho}_R^A$.

Now if the measurements of $\widehat{p}_B$ are *unrecorded* then the *conditioned variance* is

$$
\begin{aligned}
\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B} &= \sum_{x_B} \langle \Delta \widehat{p}_A^2 \rangle_{\widehat{x}_B, x_B} P(\widehat{x}_B, x_B) \\
&= \sum_{x_B} \sum_R P_R \, \langle \Delta \widehat{p}_A^2 \rangle_{\widehat{x}_B, x_B}^R \, P_B^R(\widehat{x}_B, x_B) \tag{170}
\end{aligned}
$$

which in view of inequality (50) satisfies

$$
\begin{aligned}
\langle \Delta \widehat{p}_A^2 \rangle_{\widehat{p}_B} &\geq \sum_{p_B} \sum_R P_R \, \langle \Delta \widehat{p}_A^2 \rangle^R \, P_B^R(\widehat{p}_B, p_B) \\
&= \sum_R P_R \, \langle \Delta \widehat{p}_A^2 \rangle^R \tag{171}
\end{aligned}
$$

using $\sum_{p_B} P_B^R(\widehat{p}_B, p_B) = 1$. Thus the variance for measurement of momentum $\widehat{p}_A$ conditioned on unrecorded measurements for momentum $\widehat{p}_B$ satisfies an inequality that only depends on the variances for measurements of $\widehat{p}_A$ in the possible sub-system A states $\widehat{\rho}_R^A$.

7 Appendix 2 - Inequalities

These inequalities are examples of Schwarz inequalities.

7.1 Integral Inequality

If $C(\lambda), D(\lambda)$ are real, positive functions of λ and $P(\lambda)$ is another real, positive function then we can show that

$$\int d\lambda \, P(\lambda)C(\lambda). \int d\lambda \, P(\lambda)D(\lambda) \geq \left(\int d\lambda \, P(\lambda)\sqrt{C(\lambda)D(\lambda)} \right)^2 \qquad (172)$$

To show this write $x = \int d\lambda \, P(\lambda)C(\lambda)$ and $y = \int d\lambda \, P(\lambda)D(\lambda)$. Then

$$
\begin{aligned}
xy &= \int d\lambda \, P(\lambda)C(\lambda) \int d\mu \, P(\mu)D(\mu) \\
&= \int \int d\lambda \, d\mu \, P(\lambda)P(\mu)C(\lambda)D(\mu) \\
&= \int d\lambda \, P(\lambda)^2 C(\lambda)D(\lambda) + \int \int d\lambda \, d\mu \, (1 - \delta(\lambda - \mu)) \, P(\lambda)P(\mu)C(\lambda)D(\mu)
\end{aligned}
$$

$$(173)$$

Also, write $z = \left(\int d\lambda \, P(\lambda)\sqrt{C(\lambda)D(\lambda)} \right)^2$. Then

$$
\begin{aligned}
z &= \int d\lambda \, P(\lambda)\sqrt{C(\lambda)D(\lambda)} \int d\mu \, P(\mu)\sqrt{C(\mu)D(\mu)} \\
&= \int \int d\lambda \, d\mu \, P(\lambda)P(\mu) \, \sqrt{C(\lambda)D(\lambda)}\sqrt{C(\mu)D(\mu)} \\
&= \int d\lambda \, P(\lambda)^2 C(\lambda)D(\lambda) + \int \int d\lambda \, d\mu \, (1 - \delta(\lambda - \mu)) \, P(\lambda)P(\mu) \, \sqrt{C(\lambda)D(\lambda)}\sqrt{C(\mu)D(\mu)}
\end{aligned}
$$

$$(174)$$

so that

$$
\begin{aligned}
xy - z &= \int \int d\lambda \, d\mu \, (1 - \delta(\lambda - \mu)) \, P(\lambda)P(\mu) \left(C(\lambda)D(\mu) - \sqrt{C(\lambda)D(\lambda)}\sqrt{C(\mu)D(\mu)} \right) \\
&= \frac{1}{2} \int \int d\lambda \, d\mu \, (1 - \delta(\lambda - \mu)) \, P(\lambda)P(\mu) \left(C(\lambda)D(\mu) + C(\mu)D(\lambda) - 2\sqrt{C(\lambda)D(\mu)}\sqrt{C(\mu)D(\lambda)} \right) \\
&= \frac{1}{2} \int \int d\lambda \, d\mu \, (1 - \delta(\lambda - \mu)) \, P(\lambda)P(\mu) \left(\sqrt{C(\lambda)D(\mu)} - \sqrt{C(\mu)D(\lambda)} \right)^2 \\
&\geq 0
\end{aligned}
$$

$$(175)$$

which proves the result.

For the special case where $D(\lambda) = 1$ and where $\int d\lambda \, P(\lambda) = 1$ we get the simpler result

$$\int d\lambda \, P(\lambda)C(\lambda) \geq \left(\int d\lambda \, P(\lambda)\sqrt{C(\lambda)} \right)^2 \qquad (176)$$

7.2 Sum Inequality

If C_R and D_R are real, positive quantities for various R and P_R is another real, positive quantity then we can show that

$$\sum_R P_R C_R \sum_R P_R D_R \geq \left(\sum_R P_R \sqrt{C_R D_R} \right)^2 \qquad (177)$$

To prove this write $x = \sum_R P_R C_R$ and $y = \sum_R P_R D_R$ Then

$$
\begin{aligned}
xy &= \sum_R P_R C_R \sum_S P_S D_S \\
&= \sum_R \sum_S P_R P_S C_R D_S \\
&= \sum_R P_R^2 C_R D_R + \sum_R \sum_S (1 - \delta_{RS}) P_R P_S C_R D_S \qquad (178)
\end{aligned}
$$

Also, write $z = \left(\sum_R P_R \sqrt{C_R D_R} \right)^2$. Then

$$
\begin{aligned}
z &= \left(\sum_R P_R \sqrt{C_R D_R} \right) \left(\sum_S P_S \sqrt{C_S D_S} \right) \\
&= \sum_R \sum_S P_R P_S \sqrt{C_R D_R} \sqrt{C_S D_S} \\
&= \sum_R P_R^2 C_R D_R + \sum_R \sum_S (1 - \delta_{RS}) P_R P_S \sqrt{C_R D_R} \sqrt{C_S D_S} \qquad (179)
\end{aligned}
$$

so that

$$
\begin{aligned}
xy - z &= \sum_R \sum_S P_R P_S (1 - \delta_{RS}) \left(C_R D_S - \sqrt{C_R D_R} \sqrt{C_S D_S} \right) \\
&= \frac{1}{2} \sum_R \sum_S P_R P_S (1 - \delta_{RS}) \left(C_R D_S + C_S D_R - 2 \sqrt{C_R D_S} \sqrt{C_S D_R} \right) \\
&= \frac{1}{2} \sum_R \sum_S P_S P_R (1 - \delta_{RS}) \left(\sqrt{C_R D_S} - \sqrt{C_S D_R} \right)^2 \\
&\geq 0 \qquad (180)
\end{aligned}
$$

which proves the result.

For the special case where $D_R = 1$ and where $\sum_R P_R = 1$ we get the simpler result

$$
\sum_R P_R C_R \geq \left(\sum_R P_R \sqrt{C_R} \right)^2 \qquad (181)
$$

This inequality is used in [54].

8 Appendix 3 - Spin EPR Paradox

8.1 Local Spin Operators

For two sub-systems 1 and 2 there are numerous possibilities for defining separate commuting *spin operators* for the two systems. One situation of interest is where each sub-system is associated with two *bosonic modes*, the standard annihilation operators being $\widehat{a}_1$ and $\widehat{b}_1$ for system 1 and $\widehat{a}_2$ and $\widehat{b}_2$ for system 2. The *local spin operators* for each sub-system can be defined as

$$
\begin{aligned}
\widehat{S}_x^1 &= (\widehat{b}_1^\dagger \widehat{a}_1 + \widehat{a}_1^\dagger \widehat{b}_1)/2 & \widehat{S}_y^1 &= (\widehat{b}_1^\dagger \widehat{a}_1 - \widehat{a}_1^\dagger \widehat{b}_1)/2i & \widehat{S}_z^1 &= (\widehat{b}_1^\dagger \widehat{b}_1 - \widehat{a}_1^\dagger \widehat{a}_1)/2 \\
\widehat{S}_x^2 &= (\widehat{b}_2^\dagger \widehat{a}_2 + \widehat{a}_2^\dagger \widehat{b}_2)/2 & \widehat{S}_y^2 &= (\widehat{b}_2^\dagger \widehat{a}_2 - \widehat{a}_2^\dagger \widehat{b}_2)/2i & \widehat{S}_z^2 &= (\widehat{b}_2^\dagger \widehat{b}_2 - \widehat{a}_2^\dagger \widehat{a}_2)/2
\end{aligned}
\tag{182}
$$

These satisfy the usual angular momentum commutation rules and those or the different sub-systems commute. The squares of the local vector spin operators are related to the total number operators $\widehat{N}_1 = \widehat{b}_1^\dagger \widehat{b}_1 + \widehat{a}_1^\dagger \widehat{a}_1$ and $\widehat{N}_2 = \widehat{b}_2^\dagger \widehat{b}_2 + \widehat{a}_2^\dagger \widehat{a}_2$ as $\sum_\alpha (\widehat{S}_\alpha^1)^2 = (\widehat{N}_1/2)(\widehat{N}_1/2 + 1)$ and $\sum_\alpha (\widehat{S}_\alpha^2)^2 = (\widehat{N}_2/2)(\widehat{N}_2/2 + 1)$. The *total spin operators* are

$$
\widehat{S}_\alpha = \widehat{S}_\alpha^1 + \widehat{S}_\alpha^2 \qquad \alpha = x, y, z
\tag{183}
$$

and these satisfy the usual angular momentum commutation rules.

8.2 Conditional Variances

The question is whether the conditional variances $\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}}$ for measuring $\widehat{S}_{x1}$ for sub-system 1 having measured $\widehat{S}_{x2}$ for sub-system 2, and $\left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{p}_B}$ for measuring $\widehat{S}_{y1}$ for sub-system 1 having measured $\widehat{S}_{y2}$ for sub-system 2 violate the Heisenberg Uncertainty Principle.

$$
\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} < \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle|^2
\tag{184}
$$

where the measurements on sub-system 2 are left unrecorded. If this inequality holds we have an *EPR violation*.

For *separable states* the *conditional probability* that measurement of $\widehat{S}_{x1}$ on sub-system 1 leads to eigenvalue s_{x1} given that measurement of $\widehat{S}_{x2}$ on sub-system 2 leads to eigenvalue s_{x2} is obtained from Eq.(30) as

$$
P(\widehat{S}_{x1}, s_{x1} | \widehat{S}_{x2}, s_{x2}) = \sum_R P_R\, P_1^R(\widehat{S}_{x1}, s_{x1}) P_2^R(\widehat{S}_{x2}, s_{x2}) / \sum_R P_R\, P_2^R(\widehat{S}_{x2}, s_{x2})
\tag{185}
$$

where

$$
P_1^R(\widehat{S}_{x1}, s_{x1}) = Tr_1(\widehat{\Pi}_{s_{x1}}^1 \widehat{\rho}_R^1) \qquad P_2^R(\widehat{S}_{x2}, s_{x2}) = Tr_2(\widehat{\Pi}_{s_{x2}}^2 \widehat{\rho}_R^2)
\tag{186}
$$

are the probabilities for position measurements in the separate sub-systems. The probability that measurement of $\widehat{S}_{x2}$ on sub-system 2 leads to eigenvalue s_{x2} is

$$P(\widehat{S}_{x2}, s_{x2}) = \sum_R P_R\, P_2^R(\widehat{S}_{x2}, s_{x2}) \tag{187}$$

The *mean* result for measurement of $\widehat{S}_{x1}$ for this *conditional* measurement is from Eq.(19)

$$\begin{aligned}
\left\langle \widehat{S}_{x1} \right\rangle_{\widehat{S}_{x2},s_{x2}} &= \sum_{s_{x1}} s_{x1}\, P(\widehat{S}_{x1}, s_{x1} | \widehat{S}_{x2}, s_{x2}) \\
&= \sum_R P_R \left\langle \widehat{S}_{x1} \right\rangle_R P_2^R(\widehat{S}_{x2}, s_{x2}) / P(\widehat{S}_{x2}, s_{x2})
\end{aligned} \tag{188}$$

where

$$\left\langle \widehat{S}_{x1} \right\rangle_R = \sum_{s_{x1}} s_{x1} P_1^R(\widehat{S}_{x1}, s_{x1}) \tag{189}$$

is the *mean* result for measurement of $\widehat{S}_{x1}$ when the sub-system is in state $\widehat{\rho}_R^1$.

The *conditional variance* for measurement of $\widehat{S}_{x1}$ for the conditional measurement of $\widehat{S}_{x2}$ on sub-system 2 which led to eigenvalue s_{x2} is from Eq.(20)

$$\begin{aligned}
\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2},s_{x2}} &= \sum_{s_{x1}} (s_{x1} - \left\langle \widehat{S}_{x1} \right\rangle_{\widehat{S}_{x2},s_{x2}})^2\, P(\widehat{S}_{x1}, s_{x1} | \widehat{S}_{x2}, s_{x2}) \\
&= \sum_R P_R \left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2},s_{x2}}^R P_2^R(\widehat{S}_{x2}, s_{x2}) / P(\widehat{S}_{x2}, s_{x2})
\end{aligned} \tag{190}$$

where

$$\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2},s_{x2}}^R = \sum_{s_{x1}} (s_{x1} - \left\langle \widehat{S}_{x1} \right\rangle_{\widehat{S}_{x2},s_{x2}})^2\, P_1^R(\widehat{S}_{x1}, s_{x1})$$

is a variance for measurement of $\widehat{S}_{x1}$ for when the sub-system is in state $\widehat{\rho}_R^1$ but now with the fluctuation about the mean $\left\langle \widehat{S}_{x1} \right\rangle_{\widehat{S}_{x2},s_{x2}}$ for measurements conditional on measuring $\widehat{S}_{x2}$.

However, for each sub-system state R the quantity $\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2},s_{x2}}^R$ is *minimised* if $\left\langle \widehat{S}_{x1} \right\rangle_{\widehat{S}_{x2},s_{x2}}$ is replaced by the unconditioned mean $\left\langle \widehat{S}_{x1} \right\rangle_R$ just determined from $\widehat{\rho}_R^1$. Thus we have an inequality

$$\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2},s_{x2}}^R \geq \left\langle \Delta\widehat{S}_{x1}^2 \right\rangle^R \tag{191}$$

where

$$\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle^R = \sum_{s_{x1}} (s_{x1} - \left\langle \widehat{S}_{x1} \right\rangle)^2\, P_1^R(\widehat{S}_{x1}, s_{x1}) \tag{192}$$

103

is the *normal variance* for measurement of $\widehat{S}_{x1}$ for when the sub-system is in state $\widehat{\rho}_R^1$.

Now if the measurements of $\widehat{S}_{x2}$ are *unrecorded* - as would be the case from the point of view of the experimenter on spatially well-separated sub-system 1 when measurements on this sub-system take place at the same time - then the *conditioned variance* is

$$
\begin{aligned}
\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} &= \sum_{s_{x2}} \left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}, s_{x2}} P(\widehat{S}_{x2}, s_{x2}) \\
&= \sum_{s_{x2}} \sum_{R} P_R \left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}, s_{x2}}^{R} P_2^R(\widehat{S}_{x2}, s_{x2}) \qquad (193)
\end{aligned}
$$

which in view of inequality (50) satisfies

$$
\begin{aligned}
\left\langle \Delta \widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} &\geq \sum_{s_{x2}} \sum_{R} P_R \left\langle \Delta \widehat{S}_{x1}^2 \right\rangle^{R} P_2^R(\widehat{S}_{x2}, s_{x2}) \\
&= \sum_{R} P_R \left\langle \Delta \widehat{S}_{x1}^2 \right\rangle^{R} \qquad (194)
\end{aligned}
$$

using $\sum_{s_{x2}} P_2^R(\widehat{S}_{x2}, s_{x2}) = 1$. Thus the variance for measurement of spin $\widehat{S}_{x1}$ conditioned on unrecorded measurements for spin $\widehat{S}_{x2}$ satisfies an inequality that only depends on the variances for measurements of $\widehat{S}_{x1}$ in the possible sub-system 1 states $\widehat{\rho}_R^1$.

Now exactly the same treatment can be carried out for the variance of *spin* $\widehat{S}_{y1}$ **also** conditioned on unrecorded measurements of measurements for momentum $\widehat{S}_{y2}$. Details are given in Appendix 6. We have with

$$
\begin{aligned}
\left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} &= \sum_{s_{y2}} \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}, p_{sy2}} P(\widehat{S}_{y2}, s_{y2}) \\
\left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}, s_{y2}} &= \sum_{s_{y1}} (s_{y1} - \left\langle \widehat{S}_{y1} \right\rangle_{\widehat{S}_{y2}, s_{y2}})^2 \, P(\widehat{S}_{y2}, s_{y2} | \widehat{S}_{y2}, s_{y2}) \\
\left\langle \widehat{S}_{y1} \right\rangle_{\widehat{S}_{y2}, s_{y2}} &= \sum_{s_{y1}} s_{y1} \, P(\widehat{S}_{y1}, s_{y1} | \widehat{S}_{y2}, s_{y2})
\end{aligned}
$$

the inequality

$$
\left\langle \Delta \widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} \geq \sum_{R} P_R \left\langle \Delta \widehat{S}_{y1}^2 \right\rangle^{R} \qquad (195)
$$

with

$$
\left\langle \Delta \widehat{S}_{y1}^2 \right\rangle^{R} = \sum_{s_{y1}} (s_{y1} - \left\langle \widehat{S}_{y1} \right\rangle)^2 \, P_1^R(\widehat{S}_{y1}, s_{y1}) \qquad (196)
$$

is the normal variance for measurement of $\widehat{S}_{y1}$ for when the sub-system is in state $\widehat{\rho}_R^1$.

We now multiply the two conditional variances, which it is important to note were associated with two *different conditioned states* based on two *different* measurements - spin $\widehat{S}_{x2}$ and spin $\widehat{S}_{y2}$ - carried out on sub-system 2.

$$\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} \geq \sum_R P_R \left\langle \Delta\widehat{S}_{x1}^2 \right\rangle^R \sum_S P_S \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle^S \qquad (197)$$

However, from the general inequality in Eq.(177)

$$\sum_R P_R\, C_R \sum_R P_R\, D_R \geq \left(\sum_R P_R \sqrt{C_R D_R} \right)^2 \qquad (198)$$

we then have

$$\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} \geq \left(\sum_R P_R \sqrt{\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle^R \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle^R} \right)^2 \qquad (199)$$

But we know from the HUP that for *any* given state $\widehat{\rho}_R^1$ that $\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle^R \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle^R \geq \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle^R|^2$, so for the conditioned variances associated with a *separable* state

$$
\begin{aligned}
\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}} \left\langle \Delta\widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}} &\geq \frac{1}{4} \left(\sum_R P_R \, |\left\langle \widehat{S}_{z1} \right\rangle^R| \right)^2 \\
&> \frac{1}{4} \left(\sum_R P_R \left\langle \widehat{S}_{z1} \right\rangle^R \right)^2 \\
&= \frac{1}{4} |\left\langle \widehat{S}_{z1} \right\rangle|^2 \qquad (200)
\end{aligned}
$$

showing that for a separable state the conditioned variances $\left\langle \Delta\widehat{S}_{x1}^2 \right\rangle_{\widehat{S}_{x2}}$ and $\left\langle \Delta\widehat{S}_{y1}^2 \right\rangle_{\widehat{S}_{y2}}$ *still satisfy* the HUP. It is important to note that these variances were associated with two different conditioned states based on two different measurements - spin $\widehat{S}_{x2}$ and spin $\widehat{S}_{y2}$ - carried out on sub-system 2, the results of which the observer for sub-system 1 would be unaware of. Thus if the EPR violations **as** defined in Eq.(184) are to occur then the state must be entangled.

9 Appendix 4 - Extracting Entanglement due to Symmetrisation

9.1 Two Particle Case - Bosons

The approach of Killoran et al [95] can be first applied to the simple case of $N = 2$ bosons initially in the A modes $a0$ and $a1$ and were discussed in SubSection 3.1.1. Here we present the detailed derivation of the results. The B modes $b0$ and $b1$ are initially unoccupied.

The occupied state is

$$|\Phi_A\rangle = \frac{1}{\sqrt{2}}\{|a0(1)\rangle\,|a1(2)\rangle + |a0(2)\rangle\,|a1(1)\rangle\} \tag{201}$$

in first quantisation. This is regarded by Killoran et al [95] as an entangled state for sub-systems consisting of particle 1 and particle 2. In second quantisation the occupied state $|\Phi_A\rangle$ and the unoccupied state $|\Phi_B\rangle$ are given by

$$
\begin{aligned}
|\Phi_A\rangle &= |1\rangle_{a0}\,|1\rangle_{a1} & |\Phi_B\rangle = |0\rangle_{b0}\,|0\rangle_{b1} \\
|\Phi_A\rangle &= \frac{(\widehat{a}_0^\dagger)}{\sqrt{1}}\frac{(\widehat{a}_1^\dagger)}{\sqrt{1}}\,|0\rangle_{a0}\,|0\rangle_{a1} & |\Phi_B\rangle = |0\rangle_{b0}\,|0\rangle_{b1} \tag{202}
\end{aligned}
$$

These are regarded as separable states for the A modes $a0$ and $a1$ and separable states for the B modes $b0$ and $b1$.

In second quantisation we consider the effect of the beam splitter on an *input* state

$$|\Phi_{in}\rangle = |\Phi_A\rangle \otimes |\Phi_B\rangle \tag{203}$$

The effect is to produce an *output* state given by

$$
\begin{aligned}
|\Phi_{out}\rangle &= \widehat{U}\,|\Phi_{in}\rangle \\
&= \frac{(r\widehat{b}_0^\dagger + t\widehat{a}_0^\dagger)}{\sqrt{1}}\frac{(r\widehat{b}_1^\dagger + t\widehat{a}_1^\dagger)}{\sqrt{1}}\,|0\rangle_{a0}\,|0\rangle_{a1} \otimes |0\rangle_{b0}\,|0\rangle_{b1} \\
&= r^2(|0\rangle_{a0}\,|0\rangle_{a1} \otimes |1\rangle_{b0}\,|1\rangle_{b1}) + rt(|0\rangle_{a0}\,|1\rangle_{a1} \otimes |1\rangle_{b0}\,|0\rangle_{b1} + |1\rangle_{a0}\,|0\rangle_{a1} \otimes |0\rangle_{b0}\,|1\rangle_{b1}) \\
&\quad + t^2(|1\rangle_{a0}\,|1\rangle_{a1} \otimes |0\rangle_{b0}\,|0\rangle_{b1}) \tag{204}
\end{aligned}
$$

Measurements can then be done on the output state based on projecting the state onto eigenstates for the number operators for the A and B mode-based sub-systems. The projectors $\widehat{\Pi}^A(N_A)$ for sub-system A onto eigenstates with $N_A = 0, 1, 2$ bosons are given by

$$
\begin{aligned}
\widehat{\Pi}^A(0) &= |0\rangle_{a0}\,|0\rangle_{a1}\,\langle 0|_{a0}\,\langle 0|_{a1} \\
\widehat{\Pi}^A(1) &= (|1\rangle_{a0}\,|0\rangle_{a1}\,\langle 1|_{a0}\,\langle 0|_{a1} + |0\rangle_{a0}\,|1\rangle_{a1}\,\langle 0|_{a0}\,\langle 1|_{a1}) \\
\widehat{\Pi}^A(2) &= (|2\rangle_{a0}\,|0\rangle_{a1}\,\langle 2|_{a0}\,\langle 0|_{a1} + |1\rangle_{a0}\,|1\rangle_{a1}\,\langle 1|_{a0}\,\langle 1|_{a1} + |0\rangle_{a0}\,|2\rangle_{a1}\,\langle 0|_{a0}\,\langle 2|_{a1}) \tag{205}
\end{aligned}
$$

with corresponding expressions for projectors $\widehat{\Pi}^B(N_B)$ for sub-system B.

To demonstrate entanglement extraction for particle based sub-systems with particle 1 in one sub-system, and particle 2 in the other sub-system we choose projectors corresponding to there being one particle in the A modal sub-system and one particle being in the B modal sub-system. Thus the output state is projected onto the states with $N_A = 1$ and $N_B = 1$ and we get after normalising

$$
\begin{aligned}
|\Phi_{out}(1,1)\rangle &= \mathcal{N}\left(\widehat{\Pi}^A(1) \otimes \widehat{\Pi}^B(1)\right)|\Phi_{out}\rangle \\
&= \frac{1}{\sqrt{2}}(|1\rangle_{a0}|0\rangle_{a1} \otimes |0\rangle_{b0}|1\rangle_{b1} + |0\rangle_{a0}|1\rangle_{a1} \otimes |1\rangle_{b0}|0\rangle_{b1})
\end{aligned}
$$
(206)

This is still a bipartite entangled state of the of two modal sub-systems, A and B.

If we construct a mathematical correspondence of the form

$$
\begin{aligned}
|a0(1)\rangle &\rightarrow |1\rangle_{a0}|0\rangle_{a1} & |a1(2)\rangle &\rightarrow |0\rangle_{b0}|1\rangle_{b1} \\
|a1(1)\rangle &\rightarrow |0\rangle_{a0}|1\rangle_{a1} & |a0(2)\rangle &\rightarrow |1\rangle_{b0}|0\rangle_{b1}
\end{aligned}
$$
(207)

we see that the projected output state given in (206) as a bipartite entangled state of the of two modal sub-systems, A and B, has the same *mathematical* form as the bipartite entangled state of the of two particle sub-systems containing particle 1 and particle 2. respectively.

9.2   Two Particle Case - Fermions

Here the details for the simple case of $N = 2$ fermions initially in the C modes $c0$ and $c1$ are presented, following the same approach as in the previous SubSection. The D modes $d0$ and $d1$ are initially unoccupied. Fermion modes are denoted c and d to distinguish them from bosonic modes a and b.

The occupied state is

$$
|\Phi_C\rangle = \frac{1}{\sqrt{2}}\{|c0(1)\rangle|c1(2)\rangle - |c0(2)\rangle|c1(1)\rangle\}
$$
(208)

in first quantisation. This is regarded by Killoran et al [95] as an entangled state for sub-systems consisting of particle 1 and particle 2. In second quantisation the occupied state $|\Phi_C\rangle$ and the unoccupied state $|\Phi_D\rangle$ are given by

$$
\begin{aligned}
|\Phi_C\rangle &= |1\rangle_{c0}|1\rangle_{c1} & |\Phi_D\rangle &= |0\rangle_{d0}|0\rangle_{d1} \\
|\Phi_C\rangle &= \frac{(\widehat{c}_0^\dagger)}{\sqrt{1}}\frac{(\widehat{c}_1^\dagger)}{\sqrt{1}}|0\rangle_{c0}|0\rangle_{c1} & |\Phi_D\rangle &= |0\rangle_{d0}|0\rangle_{d1}
\end{aligned}
$$
(209)

These are regarded as separable states for the C modes $c0$ and $c1$ and separable states for the D modes $d0$ and $d1$.

In second quantisation we consider the effect of the beam splitter on an *input* state

$$|\Phi_{in}\rangle = |\Phi_C\rangle \otimes |\Phi_D\rangle \tag{210}$$

The effect is to produce an *output* state given by

$$
\begin{aligned}
|\Phi_{out}\rangle &= \widehat{U}\,|\Phi_{in}\rangle \\
&= \frac{(r\widehat{d_0^\dagger} + t\widehat{c_0^\dagger})}{\sqrt{1}}\frac{(r\widehat{d_1^\dagger} + t\widehat{c_1^\dagger})}{\sqrt{1}}|0\rangle_{c0}\,|0\rangle_{c1}\otimes|0\rangle_{d0}\,|0\rangle_{d1} \\
&= r^2(|0\rangle_{c0}\,|0\rangle_{c1}\otimes|1\rangle_{d0}\,|1\rangle_{d1}) + rt(-|0\rangle_{c0}\,|1\rangle_{c1}\otimes|1\rangle_{d0}\,|0\rangle_{d1} + |1\rangle_{c0}\,|0\rangle_{c1}\otimes|0\rangle_{d0}\,|1\rangle_{d1}) \\
&\quad + t^2(|1\rangle_{c0}\,|1\rangle_{c1}\otimes|0\rangle_{d0}\,|0\rangle_{d1})
\end{aligned}
\tag{211}
$$

Note the minus sign in the second term - this is due to the fermion creation operators anti-commuting $\widehat{d_0^\dagger}\widehat{c_1^\dagger} = -\widehat{c_1^\dagger}\widehat{d_0^\dagger}$.

Measurements can then be done on the output state based on projecting the state onto eigenstates for the number operators for the C and D mode-based sub-systems. The projectors $\widehat{\Pi}^C(N_C)$ for sub-system C onto eigenstates with $N_C = 0, 1, 2$ bosons are given by

$$
\begin{aligned}
\widehat{\Pi}^C(0) &= |0\rangle_{c0}\,|0\rangle_{c1}\,\langle 0|_{c0}\,\langle 0|_{c1} \\
\widehat{\Pi}^C(1) &= (|1\rangle_{c0}\,|0\rangle_{c1}\,\langle 1|_{c0}\,\langle 0|_{c1} + |0\rangle_{c0}\,|1\rangle_{c1}\,\langle 0|_{c0}\,\langle 1|_{c1}) \\
\widehat{\Pi}^C(2) &= (|2\rangle_{c0}\,|0\rangle_{c1}\,\langle 2|_{c0}\,\langle 0|_{c1} + |1\rangle_{c0}\,|1\rangle_{c1}\,\langle 1|_{c0}\,\langle 1|_{c1} + |0\rangle_{c0}\,|2\rangle_{c1}\,\langle 0|_{c0}\,\langle 2|_{c1})
\end{aligned}
\tag{212}
$$

with corresponding expressions for projectors $\widehat{\Pi}^D(N_D)$ for sub-system D.

To demonstrate entanglement extraction for particle based sub-systems with particle 1 in one sub-system, and particle 2 in the other sub-system we choose projectors corresponding to there being one particle in the C modal sub-system and one particle being in the D modal sub-system. Thus the output state is projected onto the states with $N_C = 1$ and $N_D = 1$ and we get after normalising

$$
\begin{aligned}
|\Phi_{out}(1,1)\rangle &= \mathcal{N}\left(\widehat{\Pi}^C(1)\otimes\widehat{\Pi}^D(1)\right)|\Phi_{out}\rangle \\
&= \frac{1}{\sqrt{2}}(|1\rangle_{c0}\,|0\rangle_{c1}\otimes|0\rangle_{d0}\,|1\rangle_{d1} - |0\rangle_{c0}\,|1\rangle_{c1}\otimes|1\rangle_{d0}\,|0\rangle_{d1})
\end{aligned}
\tag{213}
$$

This is still a bipartite entangled state of the of two modal sub-systems, C and D.

If we construct a mathematical correspondence of the form

$$
\begin{aligned}
|c0(1)\rangle &\rightarrow |1\rangle_{c0}\,|0\rangle_{c1} & |c1(2)\rangle &\rightarrow |0\rangle_{d0}\,|1\rangle_{d1} \\
|c1(1)\rangle &\rightarrow |0\rangle_{c0}\,|1\rangle_{c1} & |c0(2)\rangle &\rightarrow |1\rangle_{d0}\,|0\rangle_{d1}
\end{aligned}
\tag{214}
$$

we see that the projected output state given in (213) as a bipartite entangled state of the of two modal sub-systems, C and D, has the same *mathematical* form as the bipartite entangled state of the of two particle sub-systems containing particle 1 and particle 2. respectively - even down to the correct minus sign in the second term.

108

9.3 Three Particle Case - Bosons

The key ideas in the approach by Killoran et al [95] are more fully illustrated by considering one of their specific cases, namely a quantum state with $N = 3$ identical bosons for a system with four modes - two A modes $a0$ and $a1$ and two B modes $b0$ and $b1$ - in which there are two bosons in mode $a0$ and one boson in mode $a1$. The other modes $b0$ and $b1$ are initially un-occupied. The modes $a0$ and $b0$ could be two different spatial modes for a bosonic atom in one hyperfine state, and $a1$ and $b1$ could be two different spatial modes for a bosonic atom in another hyperfine state. With particles labelled $1, 2$ and 3 the quantum state in terms of first quantisation is given by

$$
\begin{aligned}
|\Phi_A\rangle &= \frac{1}{\sqrt{12}}\{|a0(1)\rangle\,|a0(2)\rangle\,|a1(3)\rangle + |a0(1)\rangle\,|a0(3)\rangle\,|a1(2)\rangle \\
&\quad + |a0(2)\rangle\,|a0(1)\rangle\,|a1(3)\rangle + |a0(2)\rangle\,|a0(3)\rangle\,|a1(1)\rangle \\
&\quad + |a0(3)\rangle\,|a0(1)\rangle\,|a1(2)\rangle + |a0(3)\rangle\,|a0(2)\rangle\,|a1(1)\rangle\}
\end{aligned}
\tag{215}
$$

In first quantisation there is no state for the B modes, since the vacuum state is not recognised as a quantum state of anything. Following the approach of regarding labelled identical particles as sub-systems we consider a *bipartite* division of the three particle system with the first sub-system as consisting of particle 1 and 2 and the second sub-system consisting of particle 3. The same state $|\Phi_A\rangle$ can be written as

$$
\begin{aligned}
|\Phi_A\rangle &= \frac{1}{\sqrt{3}}\{|a0(1)\rangle\,|a0(2)\rangle\}\,|a1(3)\rangle \\
&\quad + \sqrt{\frac{2}{3}}\frac{1}{\sqrt{2}}\{|a0(1)\rangle\,|a1(2)\rangle + |a0(2)\rangle\,|a1(1)\rangle\}\,|a0(3)\rangle
\end{aligned}
\tag{216}
$$

In this form the state appears to be an *entangled* state for the two sub-systems. The first term (which has amplitude $1/\sqrt{3}$) represents a state for the sub-system of particle 1 and 2 with both particles in single particle state $|a0\rangle$ and a state for the sub-system of particle 3 with this particles in single particle state $|a1\rangle$. The second term (which has amplitude $\sqrt{2/3}$) represents a state for the sub-system of particles 1 and 2 with one particles in single particle state $|a0\rangle$ and the other in single particle state $|a1\rangle$, and a state for the sub-system of particle 3 with this particle in single particle state $|a0\rangle$. It is this entanglement which Killoran et al [95] wish to extract by applying a beam splitter to the state $|\Phi_A\rangle$, the *beam splitter* being associated with a unitary operator $\widehat{U}$ whose effect is to transform the single particle states $|ak\rangle$ and $|bk\rangle$ into linear combinations of each other involving *reflection* and *transmission* coefficients r, t as follows

$$
\begin{aligned}
\widehat{U}\,|ak\rangle &= r\,|bk\rangle + t\,|ak\rangle & k &= 0, 1 \\
\widehat{U}\,|bk\rangle &= t\,|bk\rangle - r\,|ak\rangle & k &= 0, 1
\end{aligned}
\tag{217}
$$

For simplicity r, t are assumed to be real with $r^2 + t^2 = 1$. The beam splitter is just assumed to couple spatial modes of the same hyperfine state. This unitary

operator applies irrespective of the particular particle occupying the one particle states.

In second quantisation the occupied state $|\Phi_A\rangle$ and the unoccupied state $|\Phi_B\rangle$ are given by

$$
\begin{aligned}
|\Phi_A\rangle &= |2\rangle_{a0} |1\rangle_{a1} \qquad |\Phi_B\rangle = |0\rangle_{b0} |0\rangle_{b1} \\
|\Phi_A\rangle &= \frac{(\widehat{a}_0^\dagger)^2}{\sqrt{2}} \frac{(\widehat{a}_1^\dagger)}{\sqrt{1}} |0\rangle_{a0} |0\rangle_{a1} \qquad |\Phi_B\rangle = |0\rangle_{b0} |0\rangle_{b1}
\end{aligned} \tag{218}
$$

where the Fock states are also written in terms of mode creation operators and vacuum states for the modes. In second quantisation it is clear that $|\Phi_A\rangle$ and $|\Phi_B\rangle$ themselves are respectively *separable* states for the A and B modes. In second quantisation the effect of the unitary operator associated with the beam splitter follows from (217) noting that $|ak\rangle \equiv \widehat{a}_k^\dagger |0\rangle$ and $|bk\rangle \equiv \widehat{b}_k^\dagger |0\rangle$ and is given by

$$
\widehat{U} \widehat{a}_k^\dagger \widehat{U}^{-1} = r\widehat{b}_k^\dagger + t\widehat{a}_k^\dagger \qquad \widehat{U} \widehat{b}_k^\dagger \widehat{U}^{-1} = t\widehat{b}_k^\dagger - r\widehat{a}_k^\dagger \qquad k = 0,1 \tag{219}
$$

In paper II we show that two mode beam splitters can indeed be described by equations analogous to (219). In second quantisation we consider the effect of the beam splitter on an *input* state

$$
|\Phi_{in}\rangle = |\Phi_A\rangle \otimes |\Phi_B\rangle \tag{220}
$$

The effect is to produce an *output* state given by

$$
\begin{aligned}
|\Phi_{out}\rangle &= \widehat{U} |\Phi_{in}\rangle \\
&= \frac{(r\widehat{b}_0^\dagger + t\widehat{a}_0^\dagger)^2}{\sqrt{2}} \frac{(r\widehat{b}_1^\dagger + t\widehat{a}_1^\dagger)}{\sqrt{1}} |0\rangle_{a0} |0\rangle_{a1} \otimes |0\rangle_{b0} |0\rangle_{b1} \\
&= r^3(|0\rangle_{a0} |0\rangle_{a1} \otimes |2\rangle_{b0} |1\rangle_{b1}) + r^2 t(|0\rangle_{a0} |1\rangle_{a1} \otimes |2\rangle_{b0} |0\rangle_{b1} + \sqrt{2}(|1\rangle_{a0} |0\rangle_{a1} \otimes |1\rangle_{b0} |1\rangle_{b1}) \\
&\quad + rt^2((|2\rangle_{a0} |0\rangle_{a1} \otimes |0\rangle_{b0} |1\rangle_{b1} + \sqrt{2}(|1\rangle_{a0} |1\rangle_{a1} \otimes |1\rangle_{b0} |0\rangle_{b1}) + t^3(|2\rangle_{a0} |1\rangle_{a1} \otimes |0\rangle_{b0} |0\rangle_{b1})
\end{aligned} \tag{221}
$$

Note this state is normalised to unity as $\langle\Phi_{out}|\Phi_{out}\rangle = (r^2 + t^2)^3 = 1$. The input state is a bipartite *separable* state of two modal sub-systems, one containing the two A modes $a0$ and $a1$ and the other the two B modes $b0$ and $b1$. The output state terms each are eigenstates of number operators $\widehat{N}_A = \widehat{a}_0^\dagger \widehat{a}_0 + \widehat{a}_1^\dagger \widehat{a}_1$ and $\widehat{N}_B = \widehat{b}_0^\dagger \widehat{b}_0 + \widehat{b}_1^\dagger \widehat{b}_1$ with eigenvalues $N_A = 0$, $N_B = 3$ for the r^3 term, $N_A = 1$, $N_B = 2$ for the $r^2 t$ term, $N_A = 2$, $N_B = 1$ for the rt^2 term, $N_A = 3$, $N_B = 0$ for the t^3 term. The same result as in (221) can also be obtained using (217) in conjunction with the first quantisation form of the input state given by (215) though the algebra is more complex. In contrast to the input state, the output state is a bipartite *entangled* state of two modal sub-systems, one containing the two A modes $a0$ and $a1$ and the other the two B modes $b0$ and $b1$. Both input and output states are states with the same total of $N = 3$ bosons.

Measurements can then be done on the output state based on projecting the state onto eigenstates for the number operators for the A and B mode-based sub-systems. The projectors $\widehat{\Pi}^A(N_A)$ for sub-system A onto eigenstates with $N_A = 0, 1, 2, 3$ bosons are given by

$$
\begin{aligned}
\widehat{\Pi}^A(0) &= |0\rangle_{a0} |0\rangle_{a1} \langle 0|_{a0} \langle 0|_{a1} \\
\widehat{\Pi}^A(1) &= (|1\rangle_{a0} |0\rangle_{a1} \langle 1|_{a0} \langle 0|_{a1} + |0\rangle_{a0} |1\rangle_{a1} \langle 0|_{a0} \langle 1|_{a1}) \\
\widehat{\Pi}^A(2) &= (|2\rangle_{a0} |0\rangle_{a1} \langle 2|_{a0} \langle 0|_{a1} + |1\rangle_{a0} |1\rangle_{a1} \langle 1|_{a0} \langle 1|_{a1} + |0\rangle_{a0} |2\rangle_{a1} \langle 0|_{a0} \langle 2|_{a1}) \\
\widehat{\Pi}^A(3) &= (|3\rangle_{a0} |0\rangle_{a1} \langle 3|_{a0} \langle 0|_{a1} + |2\rangle_{a0} |1\rangle_{a1} \langle 2|_{a0} \langle 1|_{a1} + |1\rangle_{a0} |2\rangle_{a1} \langle 1|_{a0} \langle 2|_{a1} + |0\rangle_{a0} |3\rangle_{a1} \langle 0|_{a0} \langle 3|_{a1})
\end{aligned}
$$
$$(222)$$

with corresponding expressions for projectors $\widehat{\Pi}^B(N_B)$ for sub-system B.

To demonstrate entanglement extraction for particle based sub-systems with particles 1 and 2 in one sub-system, and particle 3 in the other sub-system we choose projectors corresponding to there being two particles in the A modal sub-system and one particle being on the B modal sub-system. Thus the output state is projected onto the states with $N_A = 2$ and $N_B = 1$ and we get after normalising

$$
\begin{aligned}
|\Phi_{out}(2,1)\rangle &= \mathcal{N}\left(\widehat{\Pi}^A(2) \otimes \widehat{\Pi}^B(1)\right) |\Phi_{out}\rangle \\
&= \frac{1}{\sqrt{3}} |2\rangle_{a0} |0\rangle_{a1} \otimes |0\rangle_{b0} |1\rangle_{b1} + \sqrt{\frac{2}{3}}(|1\rangle_{a0} |1\rangle_{a1} \otimes |1\rangle_{b0} |0\rangle_{b1}
\end{aligned}
$$
$$(223)$$

This is still a bipartite entangled state of the of two modal sub-systems, A and B.

If we construct a mathematical correspondence of the form

$$
\begin{aligned}
|a0(1)\rangle |a0(2)\rangle &\to |2\rangle_{a0} |0\rangle_{a1} & |a1(3)\rangle &\to |0\rangle_{b0} |1\rangle_{b1} \\
\frac{1}{\sqrt{2}} \{|a0(1)\rangle |a1(2)\rangle + |a0(2)\rangle |a1(1)\rangle\} &\to |1\rangle_{a0} |1\rangle_{a1} & |a0(3)\rangle &\to |1\rangle_{b0} |0\rangle_{b1}
\end{aligned}
$$
$$(224)$$

we see that the projected output state given in (223) as a bipartite entangled state of the of two modal sub-systems, A and B, has the same *mathematical* form as the bipartite entangled state of the of two particle sub-systems containing particles 1 and 2 and particle 3. respectively. This type of result is proved in more general cases in [95] - here we have exhibited the key features of their approach in a particular case.

It is on this basis that Killoran et al [95] assert that the action of the beam splitter is to extract the entanglement due to symmetrisation that was present in the quantum state $|\Phi_A\rangle$ for the particle sub-systems containing particles 1 and 2 and particle 3. respectively. It is of course not their ingeneous mathematical derivation that is in dispute - it is the interpretation. From the point of view

of sub-systems being modes, not particles the action of the beam splitter is to create an entangled state of the two modal sub-systems, A and B from a state that was separable. That this entangled output state can be projected onto eigenstates of the number operators for the two modal sub-systems, A and B which have the same mathematical form as the presumed entangled initial state for the particle sub-systems containing particles 1 and 2 and particle 3. respectively is of course interesting, but it does not show that labeled identical particles can be regarded as physically accessible sub-systems. Apart from the logical issue that sub-systems must be both distinguishable from each other via physical measurements, it is noteworthy that the approach of Killoran et al [95] rested on physical processes that involved coupling modes, not identified indistinguishable particles.

10 Appendix 5 - Reference Frames and Super-Selection Rules

Several papers such as [62], [64], [53], [56], [34], [57], [58] explain the link between *reference frames* and *super-selection rules* (SSR). In this Appendix we present the key ideas involved.

10.1 Two Observers with Different Reference Frames

The first point to appreciate is that there are *two observers* - Alice and Charlie - who are involved in describing the *same quantum system*, which has been prepared via some physical process We will refer to Charlie as the *external* observer, Alice the *internal* observer. The system could be a *multi-mode* system involving identical particles, it could just be a *single mode* system or it could even be a *single particle* with or without spin. Alice and Charlie each describe quantum states in terms of their own *reference frames*, which might be a set of *coordinate axes* for the case of the spin or position states for the single particle system, or it could be a *large quantum system* with a well-defined reference *phase* in the case of multi-mode or single mode systems involving identical particles. Alice and Charlie may each choose from a set of possible reference frames - for the single particle case there are an infinite number of difference choices of coordinate axes for example, related to each other via *rotations* and/or *translations*. In *Situation A* - which *is not* associated with *SSR* - Alice and Charlie *do know* the relationship between their two reference frames (and can communicate this relationship via *classical communications*) - such as in the case of the single particle system when the relative orientation of their two different coordinate axes are known. In *Situation B* - which *is* associated with *SSR* - Alice and Charlie *do not know* the relationship between their two reference frames - such as in the multi-mode or single mode system involving identical particles when the relative phase between their two large quantum phase reference systems is not known. Alice and Charlie describe the same **system** via density operators $\widehat{\sigma}$ and $\widehat{\rho}$, and the key question is the *relationship* between these two operators in situations A and B and for various types of reference frames. In terms of the notation in [53] $\rho \rightarrow \widehat{\sigma}$ and $\widetilde{\rho} \rightarrow \widehat{\rho}$. In some situations the assumption that Alice even possesses a well-defined reference frame may be invalid, in which case it is important to realise that it is *Charlie's quantum state* which is of *most interest* for describing the system. This description may differ from what hypothetical observer Alice would regard as the quantum state.

10.2 Symmetry Groups

A particular relationship going from Alice's to Charlie's reference frame is specified by the *parameter g*, which in turn defines a *unitary transformation operator* $\widehat{T}(g)$ that acts in the system space. Particular examples will be listed below. If there was a third observer - Donald - and the relationship going from Charlie's to Donald's reference frame is specified by the parameter h, which in turn

defines a unitary operator $\widehat{T}(h)$, then if we symbolise the relationship going from Alice's to Donald's reference frame by the parameter hg, it follows that $\widehat{T}(hg) = \widehat{T}(h)\widehat{T}(g)$. This shows that the unitary operators satisfy one of the requirements to constitute a *group*, referred to generally as the *transformation* group. The other requirements are easily confirmed. The unitary operator $\widehat{T}(0) = \widehat{1}$ corresponding to the case where no change of reference frame occurs (specified by the parameter 0) exists, and satisfies the requirement that $\widehat{T}(0g) = \widehat{T}(0)\widehat{T}(g) = \widehat{T}(g0) = \widehat{T}(g)\widehat{T}(0)$. The unitary operator $\widehat{T}(g^{-1}) = \widehat{T}(g)^\dagger$ corresponding to the relationship specified as g^{-1} that converts Charlie's reference frame back to that of Alice exists, and satisfies the requirement that $\widehat{T}(0) = \widehat{T}(g^{-1})\widehat{T}(g) = \widehat{T}(g)\widehat{T}(g^{-1})$. Hence all the group properties are satisfied.

A few examples are as follows:

1. *Translation group* - single spinless particle system, with $\widehat{p}$, $\widehat{x}$.the momentum, position vector operators. Here $\underrightarrow{a}$ is a vector giving the translation of Charlie's cartesian axes reference frame from that of Alice, thus $g \equiv .\underrightarrow{a}$. The unitary translation operator is $\widehat{T}(\underrightarrow{a}) = \exp(i\widehat{p} \cdot \underrightarrow{a}/\hbar)$.

2. *Rotation group* - single particle system, with $\widehat{J}$ the angular momentum vector operators. Here $\underrightarrow{u}$ is a unit vector giving the axis and rotation angle ϕ for rotating Alice's cartesian axes reference frame into that of Charlie, thus $g \equiv .\underrightarrow{u}., \phi$. The unitary rotation operator is $\widehat{T}(\underrightarrow{u}, \phi) = \exp(i\phi\widehat{J} \cdot \underrightarrow{u}/\hbar)$.

3. *Particle number U(1) group* - single mode bosonic system, with $\widehat{a}$ the mode annihilation operator and $\widehat{N}_a = \widehat{a}^\dagger\widehat{a}$ the mode number operator. Here θ_a is the phase change Alice's to Charlie's reference frame. The unitary operator is $\widehat{T}(\theta_a) = \exp(i\widehat{N}_a\theta_a)$.

4. *Particle number U(1) group* - multi-mode bosonic system, with $\widehat{a}$ as a typical mode annihilation operator and $\widehat{N} = \sum_a \widehat{a}^\dagger\widehat{a}$ the total number operator. Here θ is the phase change from Alice's to Charlie's reference frame. The unitary operator is $\widehat{T}(\theta) = \exp(i\widehat{N}\theta)$.

In these examples the system operators $\widehat{p}$, $\widehat{J}$, $\widehat{N}_a$, $\widehat{N}$ etc are the *generators* of the respective groups. In many situations the generators commute with the Hamiltonian for the system (or more generally with the evolution operator that describes time evolution of the quantum state), in which case the group of unitary operators $\widehat{T}(g)$ is the *symmetry group*, and the generators are *conserved* physical quantities.

10.3 Relationships - Situation A

In *Situation A*, where the relationship between the reference frames for Alice and Charlie is *known* and specified by a *single* parameter g, Alice's description of the state $\widehat{\sigma}$ is related to Charlie's description $\widehat{\rho}$ for the same state via the unitary transformation

$$\widehat{\rho} = \widehat{T}(g)\,\widehat{\sigma}\,\widehat{T}(g)^{-1} \tag{225}$$

Note that this is a *passive* transformation - no change of state is involved, just the same state being described by two different observers.

As an example, consider the *spinless particle* and the *translation* group. If $\left|\underrightarrow{x}\right\rangle$ is a position eigenstate then $\widehat{T}(\underrightarrow{a})\left|\underrightarrow{x}\right\rangle = \left|\underrightarrow{x} - \underrightarrow{a}\right\rangle$. A pure quantum position eigenstate described by Alice as $\widehat{\sigma} = |\Phi\rangle\langle\Phi|$ with state vector $|\Phi\rangle = \left|\underrightarrow{x}\right\rangle$ would be described by Charlie as $\widehat{\rho} = |\Psi\rangle\langle\Psi|$ but now with $|\Psi\rangle = \left|\underrightarrow{x} - \underrightarrow{a}\right\rangle$, which is also a pure quantum position eigenstate but with eigenvalue $\underrightarrow{x} - \underrightarrow{a}$. This is as expected since Alices's cartesian axes have been translated by $\underrightarrow{a}$ to the origin of Charlie's axes without change of orientation. In the case of momentum eigenstates $\left|\underrightarrow{p}\right\rangle$ we have $\widehat{T}(\underrightarrow{a})\left|\underrightarrow{p}\right\rangle = \exp(i\,\underrightarrow{p}\cdot\underrightarrow{a}/\hbar)\left|\underrightarrow{p}\right\rangle$, so a pure quantum momentum eigenstate described by Alice with $|\Phi\rangle = \left|\underrightarrow{p}\right\rangle$ would be described by Charlie with $|\Psi\rangle = \exp(i\,\underrightarrow{p}\cdot\underrightarrow{a}/\hbar\left|\underrightarrow{p}\right\rangle$, which is also a pure momentum eigenstate with the same eigenvalue $\underrightarrow{p}$. Alice and Charlie describe the pure momentum eigernstate with the same density operator $\widehat{\rho} = \widehat{\sigma}$, the phase factor cancels.

For more general pure states, consider a quantum state described by Alice as $\widehat{\sigma} = |\Phi\rangle\langle\Phi|$ with state vector $|\Phi\rangle = \int d\underrightarrow{x}\,\phi(\underrightarrow{x})\left|\underrightarrow{x}\right\rangle$. States of this form can represent *localised* states when $\phi(\underrightarrow{x})$ is only significant in confined spatial regions, or they can represent *delocalised* states, such as momentum eigenstates $\left|\underrightarrow{p}\right\rangle$ when $\phi(\underrightarrow{x}) = (2\pi\hbar)^{-3/2}\exp(i\,\underrightarrow{p}\cdot\underrightarrow{x}/\hbar)$. We see that Charlie also describes a pure quantum state $\widehat{\rho} = |\Psi\rangle\langle\Psi|$ but now with $|\Psi\rangle = \widehat{T}(\underrightarrow{a})|\Phi\rangle = \int d\underrightarrow{x}\,\phi(\underrightarrow{x} + \underrightarrow{a})\left|\underrightarrow{x}\right\rangle = \int d\underrightarrow{x}\,\psi(\underrightarrow{x})\left|\underrightarrow{x}\right\rangle$, so the wavefunction is now $\psi(\underrightarrow{x}) = \phi(\underrightarrow{x} + \underrightarrow{a})$.

Note that if Alice's state vector was written in terms of momentum eigenstates $|\Phi\rangle = \int d\underrightarrow{p}\,\widetilde{\phi}(\underrightarrow{p})\left|\underrightarrow{p}\right\rangle$, then Charlie's state vector $|\Psi\rangle = \int d\underrightarrow{p}\,\widetilde{\psi}(\underrightarrow{p})\left|\underrightarrow{p}\right\rangle$ has a momentum wave function $\widetilde{\psi}(\underrightarrow{p}) = \exp(i\,\underrightarrow{p}\cdot\underrightarrow{a}/\hbar)\widetilde{\phi}(\underrightarrow{p})$ related to that of Alice by a phase factor. Note that a state which is a quantum superposition of momentum eigenstates as described by Alice is also described as a quantum superposition of momentum eigenstates by Charlie. A similar feature applies in all situation A cases, and is related to SSR *not* applying in situation A.

The case of the *particle* with *spin* and the *rotation* group is outlined in Ref. [64].

10.4   Relationships - Situation B

In *Situation B*, where on the other hand the relationship between frames is completely *unknown*, all possible transformations g must be given *equal weight*, and hence the relationship between Alice's and Charlie's description of the same

state becomes

$$
\begin{aligned}
\widehat{\rho} &= \int w(g) dg\, \widehat{T}(g)\, \widehat{\sigma}\, \widehat{T}(g)^{-1} \\
&= \mathcal{G}\left[\widehat{\sigma}\right]
\end{aligned}
\tag{226}
$$

where $\int w(g) dg$ is a symbolic integral over the parameter g, which includes a weight factor $w(g)$ so that $\int w(g) dg = 1$. This linear process connecting $\widehat{\sigma}$ to $\widehat{\rho}$ is the "$\mathcal{G}$- *twirling*" operation. Again, this is a paassive transformation.

It is straightforward to show that for any fixed parameter h that

$$
\widehat{T}(h)\, \widehat{\rho}\, \widehat{T}(h)^{-1} = \widehat{\rho}
\tag{227}
$$

showing that Charlie's density operator is $\mathcal{G}$ invariant under the transformation group - unlike the case for Situation A.

As an example, consider the *single mode* bosonic system and the *U(1)* group. If $|n_a\rangle$ is a Fock state then $\widehat{T}(\theta_a)\,|n_a\rangle = \exp(in_a\theta_a)\,|n_a\rangle$. Consider a pure quantum state described by Alice as the *Glauber coherent state* $\widehat{\sigma} = |\Phi\rangle\langle\Phi|$ with state vector $|\Phi(\beta)\rangle = \sum_{n_a} C(n_a, \beta)\,|n_a\rangle$, where $C(n_a, \beta) = \exp(-|\beta|^2/2)\,\beta^{n_a}/\sqrt{(n_a)!}$. It is straightforward to show that

$$
\widehat{T}(\theta_a)\,|\Phi(\beta)\rangle = |\Phi(\beta\exp(i\theta_a))\rangle
\tag{228}
$$

so that the Glauber coherent state is transformed into another Glauber coherent state, but with β changed via a phase factor to $\beta\exp(i\theta_a)$. The quantum state described by Charlie is given by

$$
\begin{aligned}
\widehat{\rho} &= \int \frac{d\theta_a}{2\pi}\,|\Phi(\beta\exp(i\theta_a))\rangle\langle\Phi(\beta\exp(i\theta_a))| \tag{229} \\
&= \int \frac{d\theta_a}{2\pi} \sum_{n_a}\sum_{m_a} C(n_a, \beta)\, C(m_a, \beta)^*\, \widehat{T}(\theta_a)\,|n_a\rangle\langle m_a|\,\widehat{T}(\theta_a)^\dagger \\
&= \sum_{n_a}\sum_{m_a} C(n_a, \beta)\, C(m_a, \beta)^*\,|n_a\rangle\langle m_a| \int \frac{d\theta_a}{2\pi}\exp(i[n_a - m_a]\theta_a) \\
&= \sum_{n_a} |C(n_a, \beta)|^2\,|n_a\rangle\langle n_a| \\
&= \sum_{n_a} \exp(-|\beta|^2)\,\frac{(|\beta|^2)^{n_a}}{(n_a)!}\,|n_a\rangle\langle n_a| \tag{230}
\end{aligned}
$$

which is a *mixed state* consisting of a *Poisson distribution* of *Fock states* with mean occupation number $\overline{n}_a = |\beta|^2$. In view of the first expression for $\widehat{\rho}$ it can also be thought of as a mixed state consisting of Glauber coherent states each with the same amplitude $|\beta| = \sqrt{\overline{n}_a}$, but with all phases $(\arg\beta + \theta_a)$ equally probable. Thus, whereas Alice describes the state as a pure state that

116

is a quantum superposition of Fock states with differing occupancy numbers, Charlie describes the same state as a mixed state involving a statistical mixture of number states. The former violates the SSR whereas the latter does not. A similar feature applies in all situation B cases, and is related to SSR applying in Situation B. Whether Alice could ever prepare such a state in the first place is controversial - see the discussion presented above in SubSections 3.2 and 3.4. However, *assuming* she could, the quantum state as described by Charlie is a mixed state.

The situation just studied relates of course to the debate [102] regarding whether the quantum state for a *single mode laser* operating well above threshold should be described by a Glauber coherent state or as a Poisson statistical mixture of photon number states. The first viewpoint (Alice) describes the state from the point of view of an internal observer with a reference frame, the second (Charlie) describes the same state from the point of view of an external observer for whose reference frame relationship to that of the internal observer is unknown. The debate is regarded by [64] as settled on the basis that both viewpoints are valid, they are just at cross purposes because they refer to descriptions of the same quantum state by two different observers.

It should not be thought however that the quantum state would always be described in such a fundamentally different manner for all Situation B cases. As an example, consider the *multi-mode bosonic* system and the *U(1)* group. Consider the pure quantum state described by Alice as the multi-mode N boson *Fock state* $\widehat{\sigma} = |\Phi\rangle \langle\Phi|$ with state vector $|\Phi(N)\rangle = |n_1 n_2...n_a...; N\rangle = \prod_a |n_1\rangle |n_2\rangle .. |n_a\rangle ...$, where $N = \sum_a n_a$. We have $\widehat{T}(\theta) |n_1 n_2...n_a...; N\rangle = \exp(iN\theta) |n_1 n_2...n_a...; N\rangle$, so that the same state would be described by Charlie as $\widehat{\rho} = |\Psi\rangle \langle\Psi|$ and with $|\Psi\rangle = |n_1 n_2...n_a...; N\rangle$. This is also a multi-mode N boson Fock state with exactly the same occupancies. The product $\exp(iN\theta) \exp(-iN\theta)$ of phase factors averages out to unity and here $\widehat{\rho} = \widehat{\sigma}$, so Alice and Charlie both describe the multi-mode Fock states in the same way. Another example for *two mode bosonic* systems and the *U(1)* group is provided by the one boson *Bell states* (the BS notation used here is non-conventional). These are entangled two mode states that Alice would describe via the state vectors $|\Phi^{\pm}\rangle = (|10\rangle \pm |01\rangle)/\sqrt{2}$. We have $\widehat{T}(\theta) |\Phi^{\pm}\rangle = \exp(i\theta) |\Phi^{\pm}\rangle$, so that the same state would be described by Charlie with $|\Psi^{\pm}\rangle = (|10\rangle \pm |01\rangle)/\sqrt{2}$. Again the product of phase factors averages to unity and $\widehat{\rho} = \widehat{\sigma}$, so Alice and Charlie both describe the quantum states as Bell states, and in the same form.

10.5 Dynamical and Measurement Considerations

Discussions of the relationship between equations governing the dynamical behaviour of Alice's and Charlie's density operators depend on whether the evolution is just governed by a Hamiltonian or whether master equations describing evolution affected by interactions with an external environment are involved. Such matters will not be treated in detail here, nor will the issue of relating

Alice's and Charlie's measurements. The latter issue is dealt with in [53].

However, in the case where Alice describes the *Hamiltonian evolution* of her density operator via the Liouville - von-Neumann equation

$$i\hbar\frac{\partial}{\partial t}\widehat{\sigma} = [\widehat{H}, \widehat{\sigma}] \tag{231}$$

where in Alice's frame the Hamiltonian is $\widehat{H}$, and where in addition the transformation group is also the *symmetry group* so that $\widehat{T}(g)\widehat{H}\widehat{T}(g)^{-1} = \widehat{H}$ for all g, it is easy to see that for both Situations A and B, Charlie's density operator will evolve via the same LVN equation

$$i\hbar\frac{\partial}{\partial t}\widehat{\rho} = [\widehat{H}, \widehat{\rho}] \tag{232}$$

Thus both Alice and Charlie will describe the same dynamical evolution, though of course the initial (and hence evolved) states may differ in the two cases.

10.6   Nature of Reference Frames

Reference frames of differing types are involved for the various transformation groups. The common feature is that they are thought of as *actual physical systems* themselves which are either macroscopic *classical* systems or macroscopic *quantum* systems in states associated with the *classical limit*. They are intended to be *essentially unaffected* by the presence of the systems for which they are acting as reference frames. In some cases relatively uncontroversial examples exist, such as for the *cartesian axes* associated with the *translation* and *rotation* groups associated with the single particle system. The physical reference system may be a large *magnet* whose magnetic field points in a well defined direction and defines a z axis, combined with an *electrostatic generator* whose electric field is in another well defined direction at right angles that defines an x axis. In other cases the existence of suitable reference frames is less clear.

In this SubSection we will describe possible phase reference frames as if they are entirely separated (or uncorrelated) with the system of interest. In terms of the treatment by Bartlett et al [64], [53] these are *non-implicated* reference frames. In the next SubSection and in the next Appendix phase reference frames that are correlated with the system of interest will be described - these are the so-called *implicated* reference frames of Bartlett et al.

For the large quantum system with a well-defined reference *phase* associated with the *U(1)* group in the case of multi-mode or single mode systems involving identical particles, the usual choice is a single mode bosonic system such as a single mode *BEC* or a *laser* with a large mean occupancy, and which is thought of as being prepared in a Glauber coherent state $|\Phi(\alpha)\rangle$ in order to provide the *phase reference frame*, the reference phase being $\arg\alpha$. Whether such a reference frame really exists is controversial. The discussion presented above in SubSections 3.2 and 3.4 raises the question of whether such a phase reference state could ever be prepared, so this choice of a physical phase reference is rather

unsatisfactory. However, from the point of view of this presentation we *assume* it does, so that - as in the previous example - Alice can describe the reference state as another coherent state. Again, whether Alice could ever prepare such a state is questionable.

Another possibility for a physical phase reference is a *macroscopic* low frequency *harmonic oscillator*, whose quantum energy eigenstates $|n\rangle$ - with $n = 0, 1, \dots, n_{max}$ and energies $n\,\hbar\omega$ can be used to construct phase eigenstates $|\theta_p\rangle$ with $p = 0, 1, \dots, n_{max}$ and $\theta_p = p \times 2\pi/(n_{max} + 1)$, and which are defined by [?]

$$|\theta_p\rangle = \frac{1}{\sqrt{n_{max} + 1}} \sum_{n=0}^{n_{max}} \exp(in\theta_p)\,|n\rangle \tag{233}$$

These states are orthonormal. The separation between the equally spaced phase angles $\Delta\theta = 2\pi/(n_{max} + 1)$ can be made very small if n_{max} is large enough. Under the effect of the harmonic oscillator Hamiltonian $\widehat{H} = \hbar\omega\widehat{N}$, where $\widehat{N}$ is the number operator, the phase state $|\theta_p\rangle$ evolves into $|\theta_p - \omega\Delta t\rangle$ during a time interval Δt, so if the time intervals are chosen so that $\omega\Delta t = 2\pi/(n_{max}+1)$, the phase angle θ_p changes into θ_{p-1}. Thus the system behaves like a backwards running *clock* [?], the phase angles θ_p defining the positions of the hands. If the clock initially has phase θ_p the probability of finding the clock to have phase θ_q after a time interval Δt is given by

$$P(\theta_q, \theta_p, \Delta t) = \frac{1}{(n_{max} + 1)^2} \frac{\sin^2((n_{max} + 1)\Delta/2)}{\sin^2(\Delta/2)} \tag{234}$$

where $\Delta = \theta_p - \theta_q - \omega\Delta t$. For times Δt such that $\omega\Delta t \ll 2\pi/(n_{max} + 1)$ the probability of the phase remaining as θ_p is close to unity. Thus if the phase state $|\theta_p\rangle$ is used as a phase reference, it will remain stable for a time Δt satisfying the last inequality. For $\Delta t \sim 100\mu s$ and $n_{max} \sim 10^4$ so that phase is defined to $\sim 10^{-3}$ radians, an oscillator frequency $\omega \sim 10^0$ s^{-1} would suffice for this phase reference standard. Such macroscopic oscillators do exist, though the process to prepare them in the phase reference *quantum* state $|\theta_p\rangle$ would be technically difficult. Whether such a system would be useful as a phase reference for optical fields or a BEC is another issue

10.7   Relational Description of Phase References

In this SubSection phase reference frames that are correlated with the system of interest will be described - these are the so-called *implicated* reference frames of Bartlett et al [64], [53].

One such approach to describing phase references in the *U(1)* group case is via the concept of *maps*. For simplicity consider a one mode system *S*, the basis vectors for which are Fock states $|m\rangle_S$, where it is sufficient to restrict $m = 0, 1, \dots, m_{max}$. The reference system R, will also be a one mode system with Fock states $|n\rangle_R$, where n is large. Product states $|m\rangle_S \otimes |n\rangle_R$ for the combined modes exist in the Hilbert space $H_S \otimes H_R$ and are eigenstates of the

various number operators, including the total number operator $\widehat{N}_T = \widehat{N}_S + \widehat{N}_R$ - where the eigenvalue is $l = m + n$. The product states may be listed via $m = 0, 1, ..., m_{\max}$ and $n = 0, 1, ...$ or $m = 0, 1, ..., m_{\max}$ and $l = m, m + 1, ...$. Here we will describe how a *coherent superpostion* of *number states*, such as a Glauber coherent state can be represented.

In the so-called *internalisation* or *quantisation* of the reference frame the product state $|m\rangle_S \otimes |n\rangle_R$ is mapped onto the product state $|m\rangle_S \otimes |n - m\rangle_R$ where $n \geq m_{\max}$. Thus

$$|m\rangle_S \otimes |n\rangle_R \rightarrow |m\rangle_S \otimes |n - m\rangle_R \tag{235}$$

Hence for a linear combination of system states given by

$$|\Phi\rangle_S = \sum_{m=0}^{m_{\max}} C_m |m\rangle_S \tag{236}$$

we have for the state $|\Phi\rangle_S \otimes |n\rangle_R$ in $H_S \otimes H_R$

$$|\Phi\rangle_S \otimes |n\rangle_R = \sum_{m=0}^{m_{\max}} C_m |m\rangle_S \otimes |n\rangle_R \rightarrow \sum_{m=0}^{m_{\max}} C_m |m\rangle_S \otimes |n - m\rangle_R = |\Psi_n\rangle_{RS} \tag{237}$$

The mapping results in an entangled state where there are n bosons distributed betweeen the two modes. This state $|\Psi_n\rangle_{RS}$ is a pure state which is compatible with the SSR and is in one-one correspondence with the original system state $|\Phi\rangle_S$. Note that to create this state the reference state $|n\rangle_R$ must have more bosons in it than $m_{\max}$. The density operator for the original pure system S state would be $\widehat{\sigma}_S = |\Phi\rangle_S \langle\Phi|_S$, and we note that this state violates the SSR. The state $|\Phi\rangle_S$ would be essentially a Glauber coherent state if $C_m = \exp(-|\alpha|^2/2)\alpha^m/(\sqrt{m!})$, with $m_{\max} \gg |\alpha|^2$. However, for the mapped state $|\Psi_n\rangle_{RS}$ the reduced density operator $\widehat{\rho}_S$ is given by

$$\begin{aligned}
\widehat{\rho}_S &= Tr_R(|\Psi_n\rangle_{RS} \langle\Psi|_{RS}) \\
&= \sum_{m=0}^{m_{\max}} |C_m|^2 |m\rangle_S \langle m|_S
\end{aligned} \tag{238}$$

This is a mixed state and is compatible with the SSR. For the Glauber coherent state $|\Phi\rangle_S$ this is the Poisson distribution of number states. Hence the original SSR violating superposition of number states for system S is mapped onto a state in the combined system for which the reduced density operator is a statistical mixture and is consistent with the SSR. $\widehat{\sigma}_S$ would correspond to Alice's description of the state, $\widehat{\rho}_S$ to Charlie's.

In the alternative so-called *externalisation* of the reference frame the mapping is between product states, and is the reverse of the previous mapping. The product state $|m\rangle_S \otimes |n\rangle_R$ is mapped onto the product state $|m\rangle_S \otimes |m + n\rangle_R$ in the Hilbert space $H_S \otimes H_R$ where the former is spanned by vectors $|m\rangle_S$ and the latter by vectors $|m + n\rangle_R$, and where $n \geq m_{\max}$. Thus

$$|m\rangle_S \otimes |n\rangle_R \rightarrow |m\rangle_S \otimes |m + n\rangle_R \tag{239}$$

The mapping of the $H_S \otimes H_R$ state $|\Psi_n\rangle_{RS}$ then is

$$
\begin{aligned}
|\Psi_n\rangle_{RS} &= \sum_{m=0}^{m_{\max}} C_m |m\rangle_S \otimes |n-m\rangle_R \\
&\rightarrow \sum_{m=0}^{m_{\max}} C_m |m\rangle_S \otimes |n\rangle_R = \left(\sum_{m=0}^{m_{\max}} C_m |m\rangle_S \right) \otimes |n\rangle_R = |\Xi_n\rangle_{RS}
\end{aligned}
\tag{240}
$$

The mapping results in a non-entangled state which is incompatible with the SSR. The state in the subspace H_S is a coherent superposition of number states, whilst that in H_R is a Fock state. The reduced density operator in H_S is $\widehat{\sigma}_S^{\#}$ given by

$$
\begin{aligned}
\widehat{\sigma}_S^{\#} &= Tr_R(|\Xi_n\rangle_{RS} \langle\Xi_n|_{RS}) \\
&= \sum_{m=0}^{m_{\max}} \sum_{k=0}^{m_{\max}} C_m C_k^* |m\rangle_S \langle k|_S
\end{aligned}
\tag{241}
$$

which is the same as $\widehat{\sigma}_S = |\Phi\rangle_S \langle\Phi|_S$ and involves coherences between different number states in contradiction to the SSR. Clearly this second mapping just reverses the first one.

Of these two treatments of phase reference frames, the internalisation version has a closer link to physics in that the pure state $|\Psi_n\rangle_{RS}$ can in principle be created and does lead to a way of creating a state that is in one-one correspondence with any SSR violating pure state $|\Phi\rangle_S$, though it is in the form of an entangled state of the S, R sub-systems rather than just S alone. This is an important point to note - the original SSR violating state does not exist as a state of a separate system, all that exists is an SSR compatible entangled state that is in one-one correspondence with it. However, the general process for creating a state such as $|\Psi_n\rangle_{RS}$ is not explained. For simple cases such as $|\Phi\rangle_S = (|0\rangle_S + |1\rangle_S)/\sqrt{2}$ the creation of the required state $|\Psi_n\rangle_{RS} = (|0\rangle_S \otimes |n\rangle_R + |1\rangle_S \otimes |n-1\rangle_R)/\sqrt{2}$, where $n \geq 1$ would seem feasible via the ejection of one boson from a BEC in a Fock state $|n\rangle_R$ into a previously unoccupied mode. .

10.8 Irreducible Matrix Representations and Super-selection Rules

If $|i\rangle$ $(i = 1, 2, ...)$ are a set of orthonormal basis vectors in the system state space, then the group of unitary operators $\widehat{T}(g)$ is represented by a group of *unitary matrices* $D(g)$

$$
\widehat{T}(g) |i\rangle = \sum_j D_{ji}(g) |j\rangle
\tag{242}
$$

with elements $D_{ji}(g)$, and such that $D(hg) = D(h)D(g)$ etc corresponding to the group properties of the operators. This is a *matrix representation* of the transformation group.

The theory of such group representations and their application to quantum systems is well established, following the pioneering work of Wigner in the 1930s. We can just use the results here. A key concept is that of *irreducible representations*. Within the system state space we can in general choose so-called irreducible sub-spaces, denoted as Γ_α of dimension d_α and spanned by new orthonormal basis vectors $|\Gamma_\alpha\lambda\rangle$ ($\lambda = 1, 2, \dots, d_\alpha$) such that

$$\widehat{T}(g)\,|\Gamma_\alpha\lambda\rangle = \sum_{\mu=1}^{d_\alpha} D^\alpha_{\mu\lambda}(g)\,|\Gamma_\alpha\mu\rangle \tag{243}$$

For each irreducible sub-space Γ_α there is *no* smaller sub-space for which the operation of all $\widehat{T}(g)$ just leads to linear combinations of vectors within that sub-space. The $d_\alpha \times d_\alpha$ matrices $D^\alpha(g)$ then form an irreducible matrix representation for the transformation group. For different α the representations are said to be *inequivalent*.

The irreducible matrices satisfy the so-called *great orthogonality theorem* [113]

$$\int w(g)dg\, D^\alpha_{\mu\lambda}(g)D^\beta_{\xi\tau}(g)^* = \frac{1}{d_\alpha}\delta_{\alpha\beta}\delta_{\mu\xi}\delta_{\lambda\tau} \tag{244}$$

The proof of this result is based on Schur's lemma.

The importance of the irreducible representations and the consequent orthogonality theorem lies in its application to Situation B cases, where we have seen that Charlie's density operator $\widehat{\rho}$ is invariant under any of the transformations $\widehat{T}(h)\,\widehat{\rho}\,\widehat{T}(h)^{-1} = \widehat{\rho}$. Suppose we represent $\widehat{\rho}$ in terms of the basis vectors $|\Gamma_\alpha\lambda\rangle$ associated with the irreducible representations

$$\widehat{\rho} = \sum_{\alpha\lambda}\sum_{\beta\tau} R^{\alpha\beta}_{\lambda\tau}\,|\Gamma_\alpha\lambda\rangle\,\langle\Gamma_\beta\tau| \tag{245}$$

where R will be a Hermitian, positive definite matrix with unit trace since it represents a density operator. Applying the transformation gives

$$\begin{aligned} \widehat{T}(h)\,\widehat{\rho}\,\widehat{T}(h)^{-1} &= \sum_{\alpha\lambda\mu}\sum_{\beta\tau\xi} R^{\alpha\beta}_{\lambda\tau}\,D^\alpha_{\mu\lambda}(h)\,|\Gamma_\alpha\mu\rangle\,\langle\Gamma_\beta\xi|\,D^\beta_{\xi\tau}(h)^* \\ &= \widehat{\rho} \end{aligned} \tag{246}$$

Averaging over h and using the great orthogonality theorem gives

$$\widehat{\rho} = \sum_\alpha\sum_\mu \left(\sum_\lambda \frac{1}{d_\alpha}R^{\alpha\alpha}_{\lambda\lambda}\right)|\Gamma_\alpha\mu\rangle\,\langle\Gamma_\alpha\mu| \tag{247}$$

This is in the form of a mixed state involving irreducible state vectors $|\Gamma_\alpha\mu\rangle$ each occuring with a probability P^α_μ given by

$$P^\alpha_\mu = \sum_\lambda \frac{1}{d_\alpha}R^{\alpha\alpha}_{\lambda\lambda} = P^\alpha \tag{248}$$

which is the same for all μ associated with a given irreducible representation Γ_α. This is clearly a positive real quantity and since

$$
\begin{aligned}
\sum_\alpha \sum_\mu P_\mu^\alpha &= \sum_\alpha \sum_\mu \sum_\lambda \frac{1}{d_\alpha} R_{\lambda\lambda}^{\alpha\alpha} = \sum_\alpha \sum_\lambda R_{\lambda\lambda}^{\alpha\alpha} \\
&= Tr\,\widehat{\rho} = 1
\end{aligned}
\tag{249}
$$

the probabilities sum to unity as required.

The final result for Charlie's density operator

$$
\widehat{\rho} = \sum_\alpha \sum_\mu P^\alpha \, |\Gamma_\alpha\mu\rangle \langle\Gamma_\alpha\mu|
\tag{250}
$$

demonstrates the presence of a *super-selection rule*. In Charlie's description of the quantum state there are *no coherences* between states $|\Gamma_\alpha\mu\rangle$ associated with differing irreducible representations of the transformation group. This represents the general form of the SSR for all transformation groups in Situation B cases.

As an example, consider the *U(1) group* and the *single mode* bosonic system. Since the Fock states satisfy $\widehat{T}(\theta_a) \, |n_a\rangle = \exp(in_a\theta_a) \, |n_a\rangle$ they form the basis for the irreducible representations of the U(1) group, the occupation number n_a specifying the irreducible representation and the 1×1 matrices $\exp(in_a\theta_a)$ being the unitary matrices. Hence Charlie will describe the quantum state as

$$
\widehat{\rho} = \sum_{n_a} P(n_a) \, |n_a\rangle \langle n_a|
\tag{251}
$$

which is a statistical mixture of Fock states with no coherences between different Fock states. This result is of the same form as in Eq.(141) and is in accord with the SSR on boson number.

As another example, consider the *U(1) group* and the *multi-mode* bosonic system. Here sums of products of Fock states

$$
|n_1 n_2 ... n_a ...; N\rangle = \prod_a |n_1\rangle |n_2\rangle .. |n_a\rangle ... \qquad N = \sum_a n_a
\tag{252}
$$

such that the total occupancy is $N = \sum_a n_a$ can be used to form irreducible representations for the transformation group in terms of linear combinations of the products with the same N. Writing these linear combinations as

$$
|\Psi_N^\mu\rangle = \sum_{\{n_1 n_2 ... n_a ...\}} C_{\{n_1 n_2 ... n_a ...\}}^{N\mu} \, |n_1 n_2 ... n_a ...; N\rangle
\tag{253}
$$

we have since $\widehat{T}(\theta) \, |n_1 n_2 ... n_a ...; N\rangle = \exp(iN\theta) \, |n_1 n_2 ... n_a ...; N\rangle$ we see that $\widehat{T}(\theta) \, |\Psi_N^\mu\rangle = \exp(iN\theta) \, |\Psi_N^\mu\rangle$ also, so the $|\Psi_N^\mu\rangle$ define the irreducible basis states. The total occupancy N specifies the irreducible representation, but here there

are many irreducible representations with the same N depending on the various μ. In this case Charlie will describe the state as

$$\widehat{\rho} = \sum_N \sum_\mu P_\mu^N \, |\Psi_N^\mu\rangle \langle\Psi_N^\mu| \tag{254}$$

which is a statistical mixture of multi-mode states $|\Psi_N^\mu\rangle$ all with the same total occupancy N. Although there are coherence terms between individual modal Fock states, there are no coherences between states with different total occupancy. This result is of the same form as in Eq.(100) and again is an example of a super-selection rule operating in terms of Charlie's description of the quantum state.

Finally, we note that in situation A where the relationship between the frames is known and there is no invariance for Charlie's density operator, we do not have SSR applying. For the *single particle* case and the *translation group* the momentum states $\left|\underset{\rightarrow}{p}\right\rangle$ define the irreducible representations, each specified by $\underset{\rightarrow}{p}$, and as we saw Charlie's description of the quantum state involved linear combinations of these irreducible basis vectors, in contradiction to the SSR.

10.9 Non-Entangled States

The essential feature of an *non-entangled* or *separable* state is that the sub-systems are considered to be *unrelated* to each other. Hence, both for Alice and Charlie there will be *separate reference frames* for each sub-system, with transformation groups - $\widehat{T}_A(g_a)$ for sub-system A, $\widehat{T}_B(g_b)$ for sub-system B, etc which relate the reference systems of Alice to those of Charlie. The transformations g_a, g_b, .. are different. The *overall* transformation operator would be of the form $\widehat{T}(g_a, g_b, ...) = \widehat{T}_A(g_a) \otimes \widehat{T}_B(g_b) \otimes ....$. Alice would describe a general non-entangled state as having a density operator

$$\widehat{\sigma} = \sum_R P_R \, \widehat{\sigma}_R^A \otimes \widehat{\sigma}_R^B \otimes \widehat{\sigma}_R^C \otimes ... \tag{255}$$

It then follows for Situation B where the reference frames for Alice and Charlie are unrelated, that Charlie would describe the same state via the density operator

$$\widehat{\rho} = \sum_R P_R \, \widehat{\rho}_R^A \otimes \widehat{\rho}_R^B \otimes \widehat{\rho}_R^C \otimes ... \tag{256}$$

where

$$\widehat{\rho}_R^C = \int w(g_c) dg_c \, \widehat{T}_C(g_c) \, \widehat{\sigma}_R^C \, \widehat{T}_C(g_c)^{-1} \qquad C = A, B, ... \tag{257}$$

Note that *separate* twirl operations are applied to the different sub-systems, as explicitly shown in the papers by Vaccaro et al [56] (see Section IIIA, Eqn. 3.3 therein) and Paterek et al [58] (see Section 6). This leads for general transformation groups to the *local group super-selection rule*, where the $\widehat{\rho}_R^C$ involve

no coherences between states associated with differing irreducible representations of the transformation group. We see that Charlie also describes a non-entangled state and with the same mixture probability P_R as for Alice. Thus non-entanglement or separability is a feature that is the *same* for both Alice and Charlie, as ought to be the case.

In the context of sub-systems consisting of *modes* (or sets of modes) occupied by *identical bosons*, the case of interest is Situation B, with each transformation group being U(1). Here the relationship between Charlie's and Alice's *phase reference* frames are unknown. Hence irrespective of Alice's description of the sub-system states $\widehat{\sigma}_R^A$, $\widehat{\sigma}_R^B$, ... we see from the previous section that Charlie will describe the separate sub-system states $\widehat{\rho}_R^A$, $\widehat{\rho}_R^B$, as statistical mixtures of number states for the separate modes (or total number states for the sets of modes in each sub-system). Thus from Charlie's point of view the separate mode density operators will satisfy the SSR. Thus we see that the introduction of reference frames and two observers - Charlie being the external one whose description of the quantum states is of primary interest - leads to the *same SSR outcome* as the simpler considerations set out in SubSections 3.2 and 3.4. Essentially the same considerations have been used in [47], [56] and the other papers to justify the *local photon number superselection rule*.

11 Appendix 6 - Super-Selection Rule Violations ?

11.1 Preparation of Coherent Superposition of an Atom and a Molecule ?

A key paper dealing with the coherent superposition of an atom and a molecule is that by Dowling et al [111], entitled "Observing a coherent superposition of an atom and a molecule". Essentially the process involves one atom A interacting with a BEC of different atoms B leading to the creation of one molecule AB, with the BEC being depleted by one B atom.

11.1.1 Hamiltonian

The Hamiltonian is given by

$$\widehat{H} = \hbar\omega_A \widehat{b}_A^\dagger \widehat{b}_A + \hbar\omega_M \widehat{b}_M^\dagger \widehat{b}_M + \hbar\omega_2 \widehat{b}_2^\dagger \widehat{b}_2 + \frac{\hbar\kappa}{2}(\widehat{b}_M^\dagger \widehat{b}_A \widehat{b}_2 + \widehat{b}_M \widehat{b}_A^\dagger \widehat{b}_2^\dagger) \qquad (258)$$

where $\widehat{b}_A, \widehat{b}_M$ and $\widehat{b}_2$ are standard bosonic annihilation operators for the atom, molecule and BEC modes respectively, ω_A, ω_M and ω_2 are the corresponding mode frequencies and κ defines the interaction strength for the process where a molecule is created or destroyed from/to an atom A and a BEC atom B. Δ is the frequency difference between the molecular state AB and the two separate states for atoms A and B – this is zero on Feshbach resonance - and is given by

$$\Delta = \omega_M - \omega_A - \omega_2 \qquad (259)$$

The Hamiltonian commutes with the total number operator $\widehat{N}_{tot}$, where

$$\widehat{N}_{tot} = 2\widehat{b}_M^\dagger \widehat{b}_M + \widehat{b}_A^\dagger \widehat{b}_A + \widehat{b}_2^\dagger \widehat{b}_2 \qquad (260)$$

where the molecule number operator is multipled by two.

11.1.2 Initial State

Initially the state of the system is given by the density operator Eqs (10) and (11) in the paper

$$\widehat{W}_{0L} = \int \frac{d\theta}{2\pi} \exp(-i\widehat{N}_{tot}\theta) \left|\Psi\right\rangle_{0L} \left\langle\Psi\right|_{0L} \exp(+i\widehat{N}_{tot}\theta) \qquad (261)$$

$$\left|\Psi\right\rangle_{0L} = \left|A\right\rangle \left|\beta\right\rangle \qquad (262)$$

where $\left|A\right\rangle$ is a state with one atom A and $\left|\beta\right\rangle$ is a Glauber coherent state for the BEC of atoms B.The super-operator acting on the pure state $\left|\Psi\right\rangle_{0L} \left\langle\Psi\right|_{0L}$ is called the *twirling operator*, the group of unitary operators $\exp(-i\widehat{N}_{tot}\theta)$ depend on a *phase* variable θ and are a unitary representation of $U(1)$, the *generator*

being $\widehat{N}_{tot}$. These operators act as a *symmetry group* for the system and leave the Hamiltonian invariant. The *initial state* is also given by

$$\widehat{W}_{0L} = \widehat{\rho}_{A-M}(0) \otimes \widehat{\rho}_2(0) \tag{263}$$

$$\widehat{\rho}_{A-M}(0) = |A\rangle \langle A| \tag{264}$$

$$\widehat{\rho}_2(0) = \int \frac{d\theta}{2\pi} \exp(-i\widehat{n}_2\theta) |\beta\rangle \langle\beta| \exp(+i\widehat{n}_2\theta) \tag{265}$$

$$= \sum_n p_n(<n>) |n\rangle \langle n| \tag{266}$$

$$= \int \frac{d\theta}{2\pi} |\beta \exp(-i\theta)\rangle \langle\beta \exp(-i\theta)| \tag{267}$$

where $\widehat{n}_2 = \widehat{b}_2^\dagger \widehat{b}_2$ is the number uperator for the BEC mode and $p_n(<n>) = \{\exp(-<n>) <n>^n /n!\}$ is a Poisson distribution, whose mean is $<n> = |\beta|^2$. Initially then there is one atom A and the BEC is in a statistical mixture of number states with a Poisson distribution, which is mathematically equivalent to a statistical mixture of Glauber coherent states $|\beta \exp(-i\theta)\rangle$ with the same amplitude $\sqrt{<n>}$ but with all phases $(\arg \beta + \theta)$ being equally weighted.

11.1.3   Implicated Reference Frame

In the paper by Dowling et al [111] the BEC is acting as an *implicated phase reference frame* (see [64], [53]). The state of the reference frame as described by Charlie is given by

$$\widehat{\rho}_{REF} = \widehat{\rho}_2(0) = \int \frac{d\theta}{2\pi} \exp(-i\widehat{n}_2\theta) |\beta\rangle \langle\beta| \exp(+i\widehat{n}_2\theta) \tag{268}$$

and from Eq. (258), there is an interaction between the reference BEC and the separate atom A and molecule M systems. However, because $<n> = |\beta|^2$ is very large, the BEC is essentially unchanged during the process, as reflected in the use of approximations in eqs (27), (28) of the paper. Another implicated phase reference frame situation, but involving a two mode reference frame is discussed in the paper by Paterek et al [58]

Overall, in terms of the discussion in Appendix 10 $\widehat{W}_{0L}$ would be *Charlie*'s description of the initial state, whereas *Alice* would describe it as $|\Psi\rangle_{0L} \langle\Psi|_{0L}$. Presumably in the paper by Dowling et al [111] what is referred to as the "state of the laboratory" be Charlie's reference frame, and what they refer to as the "internal reference frame" would refer to that of Alice. However, whether Alice could actually prepare such a state as $|\Psi\rangle_{0L} \langle\Psi|_{0L}$ is controversial - see SubSections 3.2 and 3.4, though here this is assumed to be possible.

11.1.4   Process - Alice and Charlie Descriptions

There are three stages in the process, the first being with the interaction that turns separate atoms A and B into the molecule AB turned on at Feshbach

resonance for a time $t = \pi/(2\kappa < n >)$, the second being free evolution at large Feshbach detuning Δ for a time τ leading to a phase factor $\phi = \Delta\tau$, the third being again with the interaction turned on at Feshbach resonance for a further time $t = \pi/(2\kappa < n >)$. The typical initial state $|\Psi\rangle_{0L}$ given by $|A\rangle\,|\beta\rangle$ (eq (11)) evolves into $|\Psi\rangle_{3L}$ given by (see eq. (32) of paper)

$$|\Psi\rangle_{3L} = \left(\sin(\frac{\phi}{2})\,|A\rangle - \exp(i\arg\beta)\,\cos(\frac{\phi}{2})\,|M\rangle \right) |\beta\rangle \qquad (269)$$

using approximations set out in eqs (27), (28) of the paper that depend on $< n >$ being large. Here $|M\rangle$ is a state with one molecule AB. Thus it looks like a coherent superposition of an atom state $|A\rangle$ and a molecule state $|M\rangle$ has been prepared, the atom plus molecule system being disentangled from the BEC. *Alice* would describe the final state of the system as $|\Psi\rangle_{3L}\,\langle\Psi|_{3L}$, so from her point of view a coherent superposition of an atom and a molecule has been prepared.

However, for *Charlie* the *final state* of the system is described by a density operator $\widehat{W}_{3L}$ which is reconstructed by applying the twirling operator to $|\Psi\rangle_{3L}\,\langle\Psi|_{3L}$. Noting that

$$\exp(-i\widehat{N}_{tot}\theta)\,|\Psi\rangle_{3L} = \left(\exp(-i\theta)\,\sin(\frac{\phi}{2})\,|A\rangle - \exp(-2i\theta)\,\exp(i\arg\beta)\,\cos(\frac{\phi}{2})\,|M\rangle \right) |\beta\exp(-i\theta)\rangle$$
$$(270)$$

and using

$$Tr_2(|\beta\exp(-i\theta)\rangle\,\langle\beta\exp(-i\theta)|) = \langle\beta\exp(-i\theta)|\beta\exp(-i\theta)\rangle = 1 \qquad (271)$$

we see that Charlie's final reduced density operator for the *atom-molecule system* is

$$
\begin{aligned}
\widehat{\rho}_{A-M}(3) &= Tr_2\widehat{W}_{3L} \\
&= Tr_2 \int \frac{d\theta}{2\pi}\exp(-i\widehat{N}_{tot}\theta)\,|\Psi\rangle_{3L}\,\langle\Psi|_{3L}\exp(+i\widehat{N}_{tot}\theta) \\
&= \int \frac{d\theta}{2\pi}\left(\exp(-i\theta)\,\sin(\frac{\phi}{2})\,|A\rangle - \exp(-2i\theta)\,\exp(i\arg\beta)\,\cos(\frac{\phi}{2})\,|M\rangle \right) \\
&\quad \times \left(\exp(+i\theta)\,\sin(\frac{\phi}{2})\,\langle A| - \exp(+2i\theta)\,\exp(-i\arg\beta)\,\cos(\frac{\phi}{2})\,\langle M| \right) \\
&= \sin^2(\frac{\phi}{2})\,|A\rangle\,\langle A| + \cos^2(\frac{\phi}{2})\,|M\rangle\,\langle M| \qquad (272)
\end{aligned}
$$

Thus the coherence terms like $|A\rangle\,\langle M|$ and $|M\rangle\,\langle A|$ do not appear in the final density operator when the average over θ (not β) is carried out.

For Charlie the density operator for the atom and molecule is of course a statistical mixture of a state with one atom and no molecule and a state with no atom and one molecule. The authors of [111] actually point this out in the paragraph after eq (35) where (presumably for the case $\phi = \pi/4$) it is stated "the

state is found to be . . . an incoherent mixture of an atom and a molecule.". The probabilities for detecting an atom A or a molecule AB are as in eq (33) of the paper. In terms of Charlie's description, the density operator at the end of the preparation process does *not* signify the existence of a coherent superposition of an atom and a molecule, as the title to the paper might be taken to imply. The existence of such a coherent superposition would of course be present in Alice's description, but it is Charlie's (laboratory) description that is more relevant.

11.1.5 Interference Effects Without SSR Violation

Note that *interference effects* are still present since the atom or molecule detection probabilities depend on the phase ϕ associated with the free evolution stage of the process. However, as in many other instances, the presence of coherence effects does not require the existence of *coherent superposition states* that violate the super-selection rule. The authors actually point this out in the paragraph after eq (35), where it is stated "we have clearly predicted the standard operational signature of coherence, namely Ramsey type fringes, but the coherence is not present in our mathematical description of the system." What they are referring to is Charlie's description of the final state - which indeed shows no such coherence, but the belief that coherent superposition states are needed to predict coherence effects is mistaken.

To drive this point home, the process can be treated with the initial state for the BEC being given as a Fock state $|N\rangle$. With the interaction being given as in Eq.(258) (eq (14) in the paper) the state vector is a simple linear combination of two terms

$$|\Psi(t)\rangle = A(t)|A\rangle|N\rangle + B(t)|M\rangle|N-1\rangle \tag{273}$$

This is of course an entangled state. Coupled equations for the two amplitudes $A(t)$ and $B(t)$ can easily be obtained and simple solutions obtained for stages where the Feshbach detuning is either zero or large. The state vector is continuous from one stage to the next , and the reduced density operator at the end of the three stage process for the atom plus molecule sub-system can be obtained. It is of the form

$$
\begin{aligned}
\widehat{\rho}_{A-M}(3) &= Tr_2(|\Psi(3)\rangle\langle\Psi(3)|) \\
&= \sin^2(\frac{\phi}{2})|A\rangle\langle A| + \cos^2(\frac{\phi}{2})|M\rangle\langle M|
\end{aligned}
\tag{274}
$$

which is of course a statistical mixture of a state with one atom and no molecule and a state with no atom and one molecule - and is exactly the same result as obtained in the paper by Dowling et al.[111]. Note that coherence effects in regard to the interferometric dependence on ϕ for measurements on the final state has been found without invoking either the description of the BEC via Glauber coherent states or the presence of a coherent superposition of an atomic and a molecular state. The result can easily be extended for the case where the BEC is initially in a statistical mixture of Fock states with differing N occuring

with a probability P_N. Each initial state $|A\rangle |N\rangle$ evolves as in Eq. (273). We then would have

$$
\begin{aligned}
\widehat{\rho}_{A-M}(3) &= Tr_2(\sum_N P_N \; |\Psi_N(3)\rangle \langle\Psi_N(3)|) \\
&= \sum_N P_N \left(\sin^2(\frac{\phi}{2}) \; |A\rangle \langle A| + \cos^2(\frac{\phi}{2}) \; |M\rangle \langle M| \right) \\
&= \sin^2(\frac{\phi}{2}) \; |A\rangle \langle A| + \cos^2(\frac{\phi}{2}) \; |M\rangle \langle M| \qquad (275)
\end{aligned}
$$

which is the same as before. Allowing for a statistical mixture of Fock states makes no difference to the interferometric result.

11.1.6   Conclusion

Dowling et al [111] state in their abstract that "we demonstrate that it is possible to perform a Ramsey-type interference experiment to exhibit a coherent superposition of a single atom and a diatomic molecule" . However the interferometric effects (involving the dependence on ϕ) cannot be said to exhibit the *existence* of such a coherent superposition, since the same interferometric results can be obtained *without* ever introducing such a quantum state. There is *not* a convincing case that quantum states that violate the super-selection rule forbidding the creation of coherent superpositions of Fock states with differing particle numbers can be *created*, even in Alice's reference system. The fact that an SSR violating state $|\Psi\rangle_{3L} \langle\Psi|_{3L}$ is created in Alice's reference system is not surprising, because in the process considered the initial state $|\beta\rangle$ for the BEC was assumed as a factor in Alice's initial state, and this was itself inconsistent with the SSR. Furthermore, such SSR violating states are not *needed* to describe coherence and interference effects, so that justification for their physical existence also fails.

11.2   Detection of Coherent Superposition of a Vacuum and a One-Boson State ?

Whether such super-selection rule violating states can be detected has also not been justified. For example, consider the state given by a superposition of a one boson state and the vacuum state (as discussed in [112]). Consider an interferometric process in which one mode A for a two mode BEC interferometer is initially in the state $\alpha |0\rangle + \beta |1\rangle$, and the other mode B is initially in the state $|0\rangle$ - thus $|\Psi(i)\rangle = (\alpha |0\rangle + \beta |1\rangle)_A \otimes |0\rangle_B$ in the usual occupancy number notation, where $|\alpha|^2 + |\beta|^2 = 1$. Modes A, B could refer to two different hyperfine states of a bosonic atom with non-relativistic energies $\hbar\omega_A$.and $\hbar\omega_B$, mode annihilation operators $\widehat{a}$, $\widehat{b}$. The modes are first coupled by a *beam splitter*, which could be a resonant microwave pulse that causes transitions between the two hyperfine

states and which can be described via a unitary operator $\widehat{U}_{BS}$ such that

$$
\begin{aligned}
\widehat{U}_{BS}(|1\rangle_A \otimes |0\rangle_B) &= (|1\rangle_A \otimes |0\rangle_B - i\,|0\rangle_A \otimes |1\rangle_B)/\sqrt{2} \\
\widehat{U}_{BS}(|0\rangle_A \otimes |1\rangle_B) &= (-i\,|1\rangle_A \otimes |0\rangle_B + |0\rangle_A \otimes |1\rangle_B)/\sqrt{2} \\
\widehat{U}_{BS}(|0\rangle_A \otimes |0\rangle_B) &= (|0\rangle_A \otimes |0\rangle_B).
\end{aligned}
\tag{276}
$$

After passing through the beam splitter the system is allowed to evolve freely for a time τ, the Hamiltonian being $\widehat{H}_{free} = (mc^2 + \hbar\omega_A)\widehat{a}^\dagger\widehat{a} + (mc^2 + \hbar\omega_B)\widehat{b}^\dagger\widehat{b}$ - where collisional effects have been ignored and the rest mass energy included for completeness. Following the free evolution stage, the modes are then coupled again via a beam splitter, and the probability of an atom being found in modes A, B then being measured. A straightforward treatment of the evolution shows that the final state is given by

$$
\begin{aligned}
|\Psi(f)\rangle = {} & \alpha(|0\rangle_A \otimes |0\rangle_B) \\
& + \beta \exp(-i\{mc^2/\hbar + \omega_A\}\tau) \\
& \times \left(\frac{1 - \exp(-i\Delta\tau)}{2}(|1\rangle_A \otimes |0\rangle_B) - i\frac{1 + \exp(-i\Delta\tau)}{2}(|0\rangle_A \otimes |1\rangle_B) \right)
\end{aligned}
\tag{277}
$$

where $\Delta = \omega_B - \omega_A$ is the detuning. The probabilities of finding one atom in modes A, B respectively are

$$
P_{10} = |\beta|^2 \sin^2(\Delta\tau/2) \qquad P_{01} = |\beta|^2 \cos^2(\Delta\tau/2)
\tag{278}
$$

Thus whilst coherence effects occur depending on the phase difference $\phi = \Delta\tau$ associated with the interferometric process, the overall detection probabilities only depend on the initial state via $|\beta|^2$. There is no dependence on the *relative phase* between α and β, as would be required if the superposition state $\alpha|0\rangle + \beta|1\rangle$ is to be specified from the measurement results. Exactly the same detection probabilities are obtained if the initial state is the mixed state $\widehat{\rho}(i) = |\alpha|^2(|0\rangle_A \langle 0|_A \otimes |0\rangle_B \langle 0|_B) + |\beta|^2(|1\rangle_A \langle 1|_A \otimes |0\rangle_B \langle 0|_B)$, in which the vacuum state for mode A occurs with a probability $|\alpha|^2$ and the one boson state for mode A occurs with a probability $|\beta|^2$. In this example the coherent superposition associated with the super-selection rule violating state would not be detected in the interferometric process. The paper by Dunningham et al [112] considers first a detection process that involves using a Glauber coherent state as one of the input states. Similar interference effects as in Eq. (278) are obtained. A second detection process in which the single term Glauber coherent state is replaced by a statistical mixture with all phases equally weighted in considered next, leading to the same interference effects. This again confirms that it is not necessary to invoke the existence of coherent superpositions of number states in order to demonstate interference effects.